Ejercicios Resueltos y Propuestos
Curso EYP2300

Primera Edición

Trabajo de Recopilación, Organización y Elaboración

Eduardo M. Rodríguez F.
Departamento de Estadística - Facultad de Matemáticas
Pontificia Universidad Católica de Chile

Santiago, Diciembre 2004

Prefacio

Con la intención de apoyar la labor docente que desarrolla el Departamento de Estadística de la Facultad de Matemáticas de la Pontificia Universidad Católica de Chile, se ha realizado un trabajo de recopilación y elaboración de ejercicios resueltos y propuestos, además de guías con respuestas para el curso EYP2300, donde algunos de los cuales ya fueron desarrollados en ayudantías y han sido parte de interrogaciones en semestre anteriores.

Queremos agradecer muy en especial a FONDEDOC, por haber confiado en este proyecto y habernos entregado todo su apoyo para poder ver realizada esta necesidad tanto para el Departamento de Estadística, como para todos los alumnos y alumnas que son beneficiados de los cursos de servicio que ofrece el mismo.

Este trabajo ha sido fruto de la labor que desarrollaron docentes y ayudantes que dictaron el curso entre los años 2000 y 2004.

Específicamente deseamos agradecer a los profesores

- Claudio Beltrán.
- Rolando de la Cruz.
- Eduardo Rodríguez.

Además quisiéramos agradecer el aporte de Patricia Jiménez, Romina Mesa, Ricardo Olea y Mario Tagle, tanto por el material donado, como por la revisión de este libro.

Atentamente.

Dirección
Departamento de Estadística
Facultad de Matemáticas

Santiago, Diciembre 2004

Índice General

1 Estadística Descriptiva 1

2 Probabilidad 23

3 Variable Aleatoria Discreta 45

4 Variable Aleatoria Continua 59

5 Estimación 87

6 Intervalos de Confianza y Test de Hipótesis 105

7 Test de Homogeneidad, Independencia y Bondad de Ajuste 119

8 Tablas de distribución 125

Capítulo 1

Estadística Descriptiva

1.1 Ejercicios Resueltos

PROBLEMA 1[1]

Los siguientes datos corresponden a tiempos de vida (en horas) de unas ratitas de laboratorio expuestas a un cierto veneno. Se quiere ver la efectividad de dicho veneno.

0.03	0.03	0.04	0.05	0.07	0.11	0.12	0.14	0.22	0.22
0.23	0.24	0.29	0.29	0.31	0.33	0.36	0.47	0.51	0.60
0.61	0.73	0.85	0.86	0.86	0.93	0.97	0.99	1.05	1.06
1.11	1.14	1.18	1.21	1.35	1.40	1.44	1.71	1.79	1.88
1.91	1.93	1.96	2.21	2.34	2.63	2.66	2.93	3.20	3.53

a) Construir la respectiva tabla de Frecuencias, (CON 7 INTERVALOS) calculando: marca de clase, intervalo, frecuencia absoluta, frecuencia absoluta acumulada, frecuencia relativa, frecuencia relativa acumulada.

b) Hacer el correspondiente Histograma para la frecuencia absoluta, comente las características de éste histograma.

c) Calcular la Media (Aritmética) y Mediana (Intervalar). Interpretar cual de las anteriores medidas de centralización representa mejor a la muestra. (Incluir en su comentario, lo visto en el histograma).

d) Obtenga el intervalo donde se encuentra el 40% central de la distribución.

e) ¿En que intervalo de tiempo mueren el 90% de las ratitas?

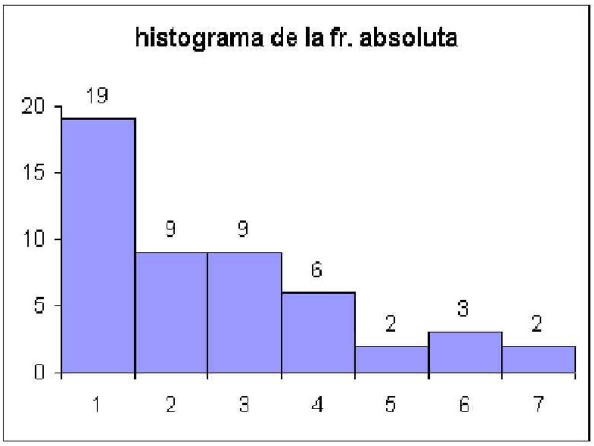
SOLUCIÓN

a) La tabla de frecuencias esta dado por:

[1] I1 segundo semestre de 2003

$\overline{x} = m_i$	L_{i-1}	L_i	n_i	N_i	f_i	F_i
0.28	0.03	0.53	19	19	0.38	0.38
0.78	0.53	1.03	9	28	0.18	0.56
1.28	1.03	1.53	9	37	0.18	0.74
1.78	1.53	2.03	6	43	0.12	0.86
2.28	2.03	2.53	2	45	0.04	0.9
2.78	2.53	3.03	3	48	0.06	0.96
3.28	3.03	3.53	2	50	0.04	1
			50			

b) En este Histograma la gran parte de los datos se encuentran en los primeros cuatro intervalos, presenta un decaimiento de ratitas muertas a medida que el tiempo de vida aumenta



c) – Media Aritmética:

$$\overline{x} = \frac{\sum_{i=1}^{7} m_i n_i}{n}$$

$$\overline{x} = \frac{0.28 \times 19 + 0.78 \times 9 + 1.28 \times 9 + 1.78 \times 6 + 2.28 \times 2 + 2.78 \times 3 + 3.28 \times 2}{50}$$

$$\overline{x} = \frac{54}{50} = 1.08$$

– Mediana:

$$Me = L_{i-1} + \frac{\frac{n}{2} - N_{i-1}}{n_i} \times a_i$$

$$Me = 0.53 + \frac{25 - 19}{9} \times 0.5 = 0.863$$

Claramente la mediana representa mejor el centro de la distribución que la media.

d) Intervalo (P_{30},P_{70})

Formula para Percentiles:

$$P_p = L_{i-1} + \frac{\frac{p\times n}{100} - N_{i-1}}{n_i} \times a_i$$

- Intervalo ocupando la formula intervalar de percentil: $(0.42, 1.41)$.
- Intervalo calculado con las respectivas posiciones: $(0.31, 1.35)$.

e) En el intervalo $(0.03, 2.53)$.

PROBLEMA 2[2]

Responda brevemente:

a) ¿Qué relación hay entre el método científico y el método estadístico?

b) De dos definiciones de tipo de muestreo.

c) Diga que se entiende por : " No depende de la unidad de medida" y señale por lo menos dos medidas que cumplan y dos medidas que no cumplan con lo antes señalado.

d) Describa en que consiste el percentil-p y haga un esquema para el cálculo en el caso de una variable.

e) ¿Qué porcentaje de la muestra está contenido en el Rango-Intercuartil?

f) Considere la **variable:Estatura**. Escriba esta variable en Escala Nominal y en Escala Ordinal.

SOLUCIÓN

a) El método estadístico es el que nos proporciona las técnicas necesarias para recolectar y analizar la información requerida, en particular la hipótesis inicial que conlleva el método científico.

b)
- Muestreo Aleatorio(m.a): indica que los elementos incluidos en esa muestra han sido seleccionados mediante algún procedimiento de sorteo o azar.
- Muestreo Aleatorio Simple(m.a.s.): cuando el procedimiento(anterior) se aplica sin restricciones sobre toda la población y cada elemento tiene igual chance de ser incluido en la muestra, hablamos de m.a.s.
- Muestreo Aleatorio Estratificado(m.a.e): cuando una población se divide(en forma natural) en sub-poblaciones o estratos(mas o menos homogéneos) se puede aprovechar esta información formando nuestra muestra en base a submuestras aleatorias simples sorteadas en cada estrato.

[2] I1 segundo semestre de 2000

c) Veamos esto a través de un ejemplo: si pensamos en medir el peso quiere decir que da lo mismo si utilizamos kilos o libras(es decir distintas unidades de medida).

caso	:	cumplan	$\rightarrowtail$	coef. variación coef. curtosis coef. asímetría
caso	:	no cumplan	$\rightarrowtail$	media varianza

d) Percentil-p (P_p)

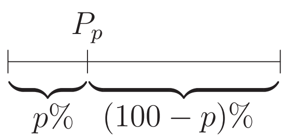


Figura 1.1:

significa que hay $p\%$ de las observaciones por debajo P_p y hay $(100-p)\%$ de las observaciones por sobre P_p

Esquema:

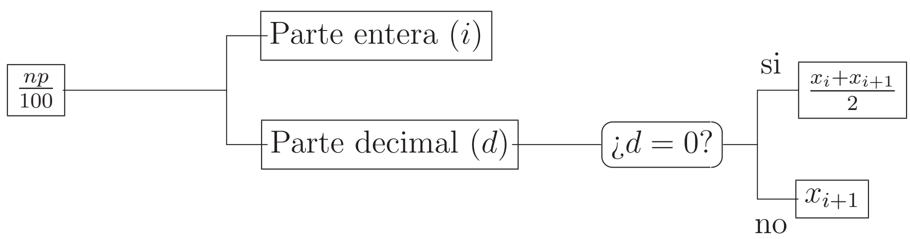


Figura 1.2:

e) El 50%. de la muestra

f) Variable: Estatura

$$\text{ESCALA NOMINAL} = \begin{cases} \text{Estatura Normal} \\ \text{Estatura No Normal} \end{cases}$$

$$\text{ESCALA ORDINAL} = \begin{cases} \text{Estatura Baja} \\ \text{Estatura Media} \\ \text{Estatura Alta} \end{cases}$$

PROBLEMA 3[3]

Los siguientes datos corresponden a las cantidades máximas de emisión diarias de óxido de azufre (en toneladas) registradas según planta de emisión, en cierta zona industrial

Cantidad de óxido(ton.)	Planta A	Planta B	Planta C
05-10	50	40	20
10-15	30	30	40
15-20	60	0	70
20-25	20	10	15
25-30	40	20	5

a) Indique la unidad de información y clasifique las variables según nivel de medición y tamaño de recorrido.

b) Entre las plantas B y C. ¿Cuál presenta mayor variabilidad relativa según la cantidad de óxido de azufre emitido?

c) ¿ Qué porcentaje de las emisiones producidas por la planta C, supera las 28 toneladas?

SOLUCIÓN

a) Unidad de información: La Planta de emisión

Variable	Según nivel de emisión	Según tamaño de recorrido
Planta de emisión	Nominal	Discreta
Cantidad de óxido	De razón	Continua

b) La siguiente tabla muestra la marca de clase (m_i) de la planta B y C y la frecuencia absoluta n_i de las mismas plantas.

[3] I1 segundo semestre de 2000

Cantidad de óxido(ton.)	m_i	n_i Planta B	n_i Planta C
05-10	7.5	40	20
10-15	12.5	30	40
15-20	17.5	0	70
20-25	22.5	10	15
25-30	27.5	20	5
		$n_B = 100$	$n_C = 150$

– Media Aritmética para la planta B y C:

$$\begin{aligned} \overline{x} &= \frac{\sum_{i=1}^{5} m_i \times n_i}{n} \\ \overline{x}_B &= \frac{40\times 7.5+...+20\times 27.5}{100} &= 14.5 \\ \overline{x}_C &= \frac{20\times 7.5+...+5\times 27.5}{150} &= 15.67 \end{aligned}$$

– Varianza para las plantas B y C:

$$\begin{aligned} S^2 &= \frac{\sum_{i=1}^{5} m_i^2 \times n_i}{n} - \overline{x}^2 \\ S_B^2 &= \frac{(7.5)^2\times 40+...+(27.5)^2\times 20}{100} - (14.5)^2 &= 61 \\ S_C^2 &= \frac{(7.5)^2\times 20+...+(27.5)^2\times 5}{150} - (15.67)^2 &= 22.37 \end{aligned}$$

Luego el coeficiente de variación (CV) esta dado por

$$\begin{aligned} CV_B &= \frac{S_B}{\overline{x}_B} &= \frac{\sqrt{61}}{14.5} &= 0.5386 \\ CV_C &= \frac{S_C}{\overline{x}_C} &= \frac{\sqrt{47.3941}}{15.67} &= 0.30 \end{aligned}$$

$\therefore$ la planta B presenta mayor variabilidad según la cantidad de óxido de azufre emitido con respecto a la planta C.

c)

$$\begin{aligned} P_p &= L_{i-1} + \frac{\frac{p\times n}{100} - N_{i-1}}{n_i} \times a_i \\ 28 &= 25 + \frac{\frac{p\times 150}{100} - 145}{5} \times 5 \\ \Rightarrow \quad p &= 98.66\% \end{aligned}$$

$\therefore 1 - p = (100 - 98.66)\% = 1.34\%$

PROBLEMA 4[4]

Lo que a continuación se presenta corresponde a la información obtenida de hoteles Internacionales. Las variables que fueron examinadas son:

TH	:	Tipo de Hotel(Categorías 1 y 2)
NH	:	Número de habitaciones
PTA	:	Precio temporada alta
PTB	:	Precio temporada baja
SC	:	Servicio cena (1:=Sí,0:=No)
TP	:	Si el cuenta con piscina(1:=Sí,0:=No)

En base a la siguiente información, responda las preguntas que al final se señalan:

Resumen de la Información

Tipo de hotel

TH	Frecuencia	Porcentaje	Frecuencia Acumulada	Porcentaje Acumulado
1	14	43.75	14	43.75
2	18	56.25	32	100

Servicio cena

TH	Frecuencia	Porcentaje	Frecuencia Acumulada	Porcentaje Acumulado
1	12	37.5	12	37.5
2	20	62.5	32	100

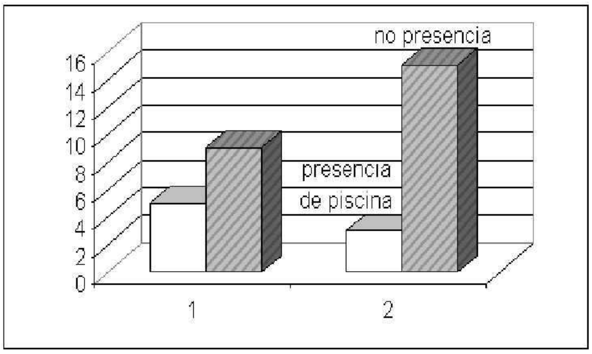


Figura 1.3: Gráfico según tipo de Hotel en relación a la Presencia o no de piscinas

Variable	N	Desv.Estándar	Suma	Mínimo	Máximo	Q1	Q2	Q3	Moda
NH	32	¿?	515	8	40	¿?	¿?	¿?	¿?
PTA	32	30.1395	3208	64	169	—	—	—	—
PTB	32	9.99921	1868	39	94	52	59	61	59

[4]I1 segundo semestre de 2000 y 2001

Diagrama de tallo y dos hojas para la variable: NH

4	0
3	
3	4
2	5 5 5
2	0 1 1 1
1	5 5 6 7 9
1	0 0 0 1 1 1 2 2 2 3 3 4 4 4
0	8 8 9 9

Preguntas:

a) Clasifique las variables según nivel de medición y tamaño de recorrido.

b) Indique la medida de tendencia central más adecuada para la variable PTB. Justifique.

c) ¿Qué medidas de posición tienen en común las variables TH,SC,TP? Justifique.

d) Resuma en una tabla Bivariada, las variables TH y TP.

e) Para las variables PTA y PTB.¿Cuál de ellas presenta menor variabilidad?

f) Para la variable NH, llene los signos ¿? con la información que se da en el diagrama de tallo y hoja.

SOLUCIÓN

a) Unidad de información: La planta de Emisión.

Variables	Según nivel de emisión	Según tamaño de recorrido
TH	Nominal	Discreta
NH	Razón	Discreta
PTA	Razón	Contínua
PTB	Razón	Contínua
SC	Nominal	Discreta(Dicotómica)
TP	Nominal	Discreta(Dicotómica)

b) Variable:PTB

$$\begin{array}{lclcl} \text{Moda} & = & 59 & & \\ \text{Mediana} & = & Q_2 & = & 59 \\ \text{Media} & = & \frac{1868}{32} & = & 58.375 \end{array}$$

Como "hay simetría"la medida de tendencia central mas adecuada es el PROMEDIO (Media).

c) La Moda, por ser del tipo Nominal.

d) Tabla Bivariada de las variables TH y TP.

		TH		
		1	2	Total
TP	0	5	3	8
	1	9	15	24
Total		14	18	32

e) El coeficiente de variación (CV) para las variables PTA y PTB son:

$$CV_{PTA} = \frac{30.1395}{(3208/32)} = 0.30064$$

$$CV_{PTB} = \frac{9.99921}{(1868/32)} = 0.17129$$

$\therefore$ la variable PTB presenta menor variabilidad con respecto a la variable PTA.

f)
- Desviación Estandar:7.6474 (n=32).
- Moda: 10, 11, 12, 14, 21, 25.
- Los Cuartiles

$$Q1 = \frac{X_8+X_9}{2} = \frac{11+11}{2} = 11$$

$$Q2 = \frac{X_{16}+X_{17}}{2} = \frac{11+11}{2} = 14$$

$$Q3 = \frac{X_{24}+X_{25}}{2} = \frac{11+11}{2} = 20.5$$

PROBLEMA 5[5]

Responda brevemente:

a) Señale las etapas en la aplicación del método científico.

b) Diga que diferencia hay entre población objetivo y población muestreada.

c) Diga cuando una variable es del tipo discreta y cuando es del tipo continua.

d) Si una variable es de nivel de medición nominal, entonces la medida de tendencia central más adecuada es la mediana. Justifique.

e) Describa en que consiste el percentil-p y haga un esquema para el cálculo en el caso de una variable.

[5] I1 recuperativa,segundo semestre de 2000

f) Considere la **variable:PESO**. Escriba esta variable en Escala Nominal y en Escala Ordinal. Además escriba esta variable según tamaño de recorrido en forma dicotómica.

SOLUCIÓN

a)
- Detección y Enunciado del Problema.
- Formulación de una hipótesis.
- Deducción de una consecuencia verificable.
- Verificación de la consecuencia.
- Conclusión.

b)

Población Objetivo	:	A quien o quienes va dirigido el estudio(Universo).
Población Muestreada	:	Una fracción de este universo a estudiar.

c)

Discreta	:	Dominio finito o infinito numerable.
Continua	:	Dominio es un intervalo en los reales.

d) Falso, Nominal$\rightarrow$ asigna nombre(no categoriza)luego no puedo calcular la mediana.

e) Percentil-p (P_p)

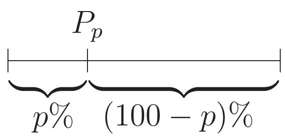


Figura 1.4:

significa que hay $p\%$ de las observaciones por debajo P_p y hay $(100-p)\%$ de las observaciones por sobre P_p.

Esquema:Ver figura 1.5

f) Variable: Peso

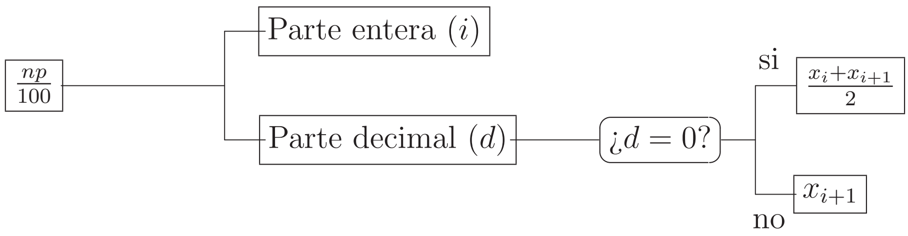


Figura 1.5:

$$\text{ESCALA NOMINAL} = \begin{cases} \text{Peso Normal} \\ \text{Peso No Normal} \end{cases}$$

$$\text{ESCALA ORDINAL} = \begin{cases} \text{Peso Bajo} \\ \text{Peso Medio} \\ \text{Peso Alto} \end{cases}$$

$$\text{DICOTÓMICA} = \begin{cases} 1 \text{ Sobre Peso} \\ 0 \text{ Bajo Peso} \end{cases}$$

PROBLEMA 6[6]

En un proceso de destilación químico, el porcentaje (Y) de pureza de oxígeno producido está relacionado con el porcentaje (X) de hidrocarburo, presente en el condensador principal de la unidad de destilación.
Se efectuaron 55 mediciones, en las cuales se observaron conjuntamente las variables X e Y, cuyos resultados se incluyen en la siguiente tabla:

	Nivel de pureza del Oxígeno (%)			
Nivel de Hidrocarburo(%)	$87-90$	$90-93$	$93-96$	$96-100$
$0.87-1.07$	10	5	0	0
$1.07-1.27$	5	12	2	1
$1.27-1.47$	1	4	9	2
$1.47-1.67$	0	1	2	1

a) ¿En qué porcentaje de las mediciones se observa un nivel de hidrocarburo superior a 1.2 % en el condensador principal, cuando en nivel de pureza de oxígeno es por lo menos 90 %?

[6] TAV 2004

b) Calcule el porcentaje de variabilidad del nivel de pureza del oxígeno para los casos en que se observa en el condensador principal un nivel de hidrocarburo inferior a 1.27 % .

SOLUCIÓN

a)

Nivel de Hidrocarburo(%)	n	N
0.97-1.07	5	5
1.07-1.27	15	20
1.27-1.47	15	35
1.47-1.67	4	39

$$
\begin{aligned}
P_p &= L_{i-1} + \frac{\frac{p\times n}{100} - N_{i-1}}{n_i} \times a_i \\
1.2 &= 1.07 + \frac{\frac{p\times 39}{100} - 5}{15} \times 0.2 \\
\Rightarrow \quad p &= 37.82\%
\end{aligned}
$$

$\therefore 1 - p = (100 - 37.82)\% = 62.18\%$

b)

Nivel de Pureza	n	m_i
87-90	15	88.5
90-93	17	91.5
93-96	2	94.5
96-100	1	98

$$
\begin{aligned}
\overline{x} &= \frac{15 \times 88.51 + 17 \times 91.5 + 2 \times 94.5 + 1 \times 98}{35} \\
&= 90.5714
\end{aligned}
$$

$$
\begin{aligned}
S^2 &= \frac{\sum_{i=1}^{5} m_i^2 \times n_i}{n} - \overline{x}^2 \\
&= 4.7215
\end{aligned}
$$

$\therefore C.V = \frac{2.729}{90.5714} \approx 0.024$

PROBLEMA 7[7]

Un fabricante de bicicletas de competición, ha reunido información sobre los diferentes sillines existentes en el mercado, respecto de alguna de las variables de interés

Marca : A, B, C.
Material : Acero(A); Aluminio(AL); Cromoly(C).

Peso en gr.
Precio en U.M.(unidades monetarias)

Long: Longitud

Si Long $\in [23, 26]$, entonces longitud=1
Si Long $\in [23, 26]$, entonces longitud=2
Si Long $\in [23, 26]$, entonces longitud=3

En base a la siguiente información, responda las preguntas que al final se señalan:

Resumen de la Información

Variable: Precio

Material: A			Material: AL			Material: C		
N	Media	Varianza	N	Media	Varianza	N	Media	Varianza
13	4282	4254805.69	25	1687.50	191406.25	15	6794.50	37711420.50

Variable: Peso

N	Media	Desv. Estándar	Coef. Variación	Mediana	Suma	Kurtosis	Moda
53	307.92	93.18	30.26	280	16320	1.31	220

Percentiles

100%	95%	90%	75%	25%
590	524	462	324	246

Tabla de Frecuencias : Marca vs. Longitud

	A	B	C	Total
1	6	2	2	10
2	1	15	4	20
3	8	13	2	23
Total	15	30	8	53

[7] I1 recuperativa, segundo semestre de 2000

Preguntas:

a) Clasifique las variables según nivel de medición y tamaño de recorrido.

b) Construya un gráfico Box-Plot que muestra la distribución de los sillines según peso, ubique los valores correspondientes en el gráfico e indique las medidas de posición y dispersión más adecuadas. Justifique su respuesta.

c) Compare la variabilidad (coef.variación) del precio de todos los sillines con la variabilidad del precio de los sillines Cromoly.(Varianza(precio de todos los sillines)=8157632.44).

d) Construya un gráfico que muestre la distribución de los sillines marca C, según longitud.

SOLUCIÓN

a)

Variables	Tamaño Recorrido	Nivel de medición
Marca	Discreta	Nominal
Material	Discreta	Nominal
Peso	Continua	Razón
Precio	Continua	Razón
Long	Discreta	Ordinal

b) Gráfico Boxplot

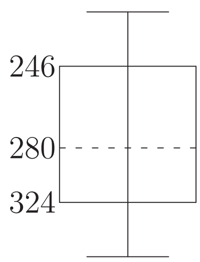


Figura 1.6:

$$\begin{array}{lcl} \text{Moda} & = & 220 \\ \text{Media} & = & 307.92 \quad \Rightarrow \quad \text{Asimetría} \\ \text{Mediana} & = & 280 \end{array}$$

Medida de Posición: Me=280 gr.
Medida de Dispersión: Q3-Q1= 324-246=78 gr.

c) Los coeficientes de variación son:

$$C.V_C = \frac{\sqrt{3711420.50}}{6794.50} = 0.2835$$
$$C.V_T = \frac{S_T}{\overline{x}_T}$$

donde

$$\overline{x}_T = \frac{4282 \times 13 + 1687.50 \times 25 + 6794.50 \times 15}{53}$$
$$= 3769.264151$$

Luego $C.V_T = \frac{2856.156935}{3769.264151} = 0.7577$ y por lo tanto $C.V_C < C.V_T$ lo que significa que los sillines marca Cromoly son más homogéneos.

d)

Long	n_c	d_i
23-26	2	2/3
27-30	4	4/3
31-30	2	2/9

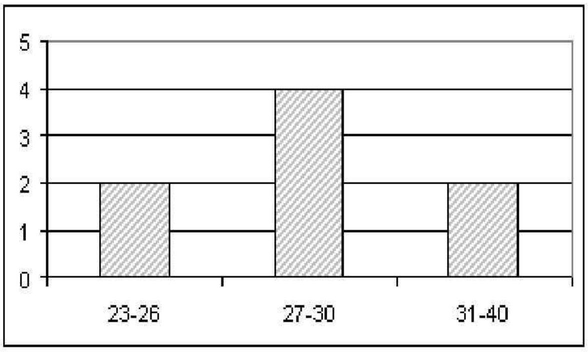


Figura 1.7:

PROBLEMA 8[8]

La distribución adjunta son las frecuencias en mediciones de resistencia a la fractura (en MPa) para barras de cerámicas quemadas.

Clase	81-83	83-85	85-87	87-89	89-91	91-93	93-95	95-97	97-99
Frecuencia	6	7	17	30	43	28	22	13	3

a) Trace un histograma basando en frecuencias relativas y comente sus propiedades interesantes.

b) ¿Qué proporción de las observaciones son cuando menos 85?¿Menores que 95?

c) Aproximadamente,¿qué proporción de las mediciones fueron menores que 90?

d) Calcule el promedio e interprete su resultado.

SOLUCIÓN

a) Propiedades: unimodal, un poco de asimetría para valores pequeños(asimetría negativa), etc.

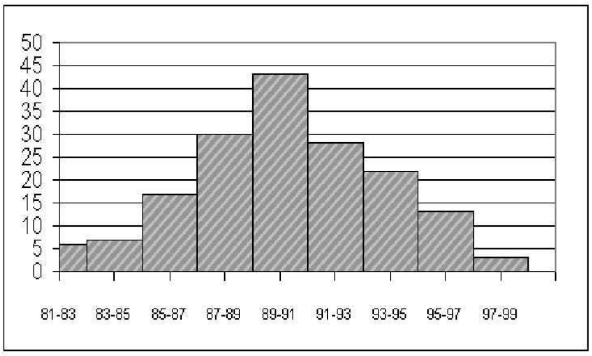


Figura 1.8:

b)

$$\begin{aligned}\text{Proporción cuando menos 85} &= \frac{17+30+43+28+22+13+3}{169}\\ &= \frac{156}{169}\\ &= 0.923\end{aligned}$$

[8]TAV 2004

ó $1 - \frac{6+7}{169} = 0.923$

$$\begin{aligned}
\text{Proporción menores } 95 &= \frac{6+7+17+30+43+28+22}{169} \\
&= \frac{153}{169} \\
&= 0.905
\end{aligned}$$

ó $1 - \frac{6+7}{169} = 0.923$

c) Aproximadamente $\frac{81.5}{169} = 0.482$

d)

$$\begin{aligned}
\overline{x} &= \frac{\sum_{i=1}^{9} m_i \times n_i}{n} \\
&= \frac{82 \times 6 + 84 \times 7 + \ldots + 98 \times 3}{169} \\
&= 90.17
\end{aligned}$$

La resistencia promedio a la fractura para barras de cerámica quemadas es de 90.17.

1.2 Ejercicios Propuestos

1. [9] La carga de fuego (Mj/m^2) es la energía calorífica que se puede liberar por metro cuadrado de área de piso por combustión del contenido y de la estructura misma de un recinto. El artículo "Fire Loads in Office Building"(*J. of Structural Engr.*, 1997, 365-368) presenta los siguientes porcentajes acumulados, leídos de una gráfica, de cargas de fuego en una muestra de 388 recintos:

Valor	**0**	**150**	**300**	**450**	**600**	**750**	**900**
% acumulado	0	19.3	37.6	62.7	77.5	87.2	93.8
Valor	**1050**	**1200**	**1350**	**1500**	**1650**	**1800**	**1950**
% acumulado	95.7	98.6	99.1	99.5	99.6	99.8	100.0

   a) Trace un histograma de frecuencia relativa y comente sus características interesantes.

   (b) ¿Qué proporción de cargas de fuego son menores que 600? ¿Por lo menos 1200?

   (c) ¿Qué proporción de cargas están entre 600 y 1200?

2. Los siguientes datos corresponden al peso recién nacidos en un hospital durante una semana.

Nº	Sexo	Peso (g)		Nº	Sexo	Peso (g)
1	M	3700		11	F	3400
2	M	2600		12	F	2700
3	F	3100		13	M	3800
4	F	3300		14	F	3500
5	F	3500		15	M	4200
6	F	3400		16	M	3300
7	F	3200		17	M	2800
8	M	2600		18	F	3600
9	M	3500		19	F	3200
10	M	2500		20	M	3500

   a) Elegir una muestra diferenciando por sexo, (muestreo aleatorio estratificado) 5 hombres y 5 mujeres utilizando muestreo sistemático. (Tomando por primer sujeto el uno).

   b) Calcular el promedio del peso por grupo.¿ Se puede decir que el sexo influye en el peso de los recién nacidos?

3. La siguiente tabla representa el número de personas activas dentro de una muestra de 50 familias:.

[9] I1 segundo semestre de 2002

Personas Activas	Fr. Absoluta N^o Familias n_i	Frec. relativa f_i	Fr. abs. acum. N_i	Fr. rel. acum. F_i
1	16			
2	20			
3	9			
4	5			
	50			

a) Complete la tabla de frecuencia.

b) ¿ Cuántas familias tiene a lo mas 2 personas activas?.

c) ¿ Qué porcentaje de familias tiene solamente una persona activa?.

4. Al comenzar el curso se paso una encuesta a los alumnos del primer curso de un colegio, preguntándoles, entre otras cosas, por el número de hermanos que tenía, obteniéndose los siguientes resultados:

3	3	2	8	2	3	4	3	3	3	5
1	3	3	2	3	2	4	3	3	2	4
3	4	3	2	4	3	2	4	2	2	3
4										

A: Represente este conjunto de datos con un diagrama de barras.

B: Calcule media, moda y mediana.

5. Al comenzar el curso se paso una encuesta a los alumnos del primer curso de un colegio, preguntándoles, entre otras cosas, por el número de hermanos que tenía, obteniéndose los siguientes resultados:

60	56	54	48	99	65	58	55	74	52	53	58	67	62	65
76	85	92	66	62	73	66	59	57	54	53	58	57	55	60
65	65	74	55	73	97	82	80	64	70	101	72	96	73	55
59	67	49	90	58	63	96	100	70	53	67	60	54		

Obtenga:

A: La distribución de frecuencias agrupando por intervalos.

B: La mediana de la distribución.

C: La media de la distribución, indicando su nivel de representatividad.

6. En el Consejo de Apuestas del Estado se han ido anotando, durante una temporada, el número de premiados de quinielas según la cantidad de aciertos, obteniéndose la siguiente tabla:

N^o de aciertos	11	12	13	14	15
N^o de personas (miles)	52	820	572	215	41

Calcule:

A: El número medio de aciertos por temporada.

B: La mediana, la moda y los cuartiles de la distribución.

7. En un puerto se controla diariamente la entrada de pesqueros según su tonelaje, resultando para un cierto día los siguientes datos:

Peso (Tm.)	0-25	25-50	50-70	70-100	100-500
N^o de barcos	5	17	30	25	3

Se pide:

A: El peso medio de los barcos que entran en el puerto diariamente, indicando la representatividad de dicha medida.

B: El intervalo donde se encuentra el 60% central de la distribución.

8. Los datos siguientes que aparecen a continuación corresponden al número de tornillos defectuosos por caja en una muestra de 90 cajas de un lote llegado a una ferretería en Marzo de 1997:

6	4	3	5	2	5	2	4	7	10
7	4	5	8	5	10	11	6	7	6
8	8	5	6	3	8	8	9	4	6
4	11	4	4	4	9	8	4	5	10
6	3	2	8	4	3	4	5	8	6
7	6	3	5	8	5	3	6	3	12
7	6	3	5	6	5	4	2	7	8
5	7	3	5	7	5	7	3	6	4
6	4	4	4	5	5	2	3	3	4

Obtenga:

A: La distribución de frecuencias.

B: La media, mediana, moda de la distribución.

C: Hacer el respectivo Histograma y gráfico de Polígonos.

9. Los datos siguientes representan en centímetros las longitudes de 50 artículos producidos por una máquina.

4,15	4,8	5,15	5,33	5,57	5,74	6,02	6,66	6,98	7,3
4,27	4,86	5,27	5,39	5,63	5,86	6,04	6,66	7,1	7,38
4,62	4,92	5,27	5,45	5,63	5,86	6,1	6,75	7,14	7,54
4,68	4,98	5,33	5,51	5,63	5,86	6,33	6,92	7,22	7,7
4,68	5,15	5,33	5,51	5,63	6,02	6,66	6,98	7,22	7,72

Obtenga:

A: Construya una tabla de frecuencias para los datos, con intervalos.

B: Construya un histograma y polígono de frecuencias para la tabla construida en a).

C: Si el 25% de los artículos de menor longitud y el 10% de los artículos de mayor longitud son considerados defectuosos por el Dpto. de control de calidad. Indique entre que longitud los artículos serán considerados aceptables.

10. Los valores de la presión sanguínea se reportan a veces a los 5 mm Hg más cercanos (100, 105, 110, etc). Suponga que los valores reales de presión sanguínea para nueve individuos seleccionados al azar son:
118.6 127.4 138.4 130.0 113.7 122.0 108.3 131.5 133.2

    a) ¿Cuál es la mediana de los valores *reportados* de la presión sanguínea?

    b) Suponga que la presión del segundo individuo es de 127.6 en lugar de 127.4 (un pequeño cambio en un sólo valor). ¿Cómo afecta esto a la mediana de los valores reportados?. ¿Qué dice esto sobre la sensibilidad de la mediana para redondear o agrupar los datos?

11. El artículo *Oxygen Consumption During Fire Suppression: Error of Heart Rate Estimation* presentó los datos siguientes sobre consumo de oxígeno en ml/Kg/min, para una muestra de diez bomberos que hicieron un simulacro de combate de incendio:
29.5 49.3 30.6 28.2 28.0 26.3 33.9 29.4 23.5 31.6

    Calcule lo siguiente:

    a) El intervalo de la muestra.

    b) La varianza muestral S^2 de la definición (es decir, calcular primeramente desviaciones con respecto a la media y luego elevarlas al cuadrado, etc.).

    c) La desviación estándar muestral.

12. Un estudio de la relación entre la edad y varias funciones visuales, por ejemplo, agudeza y percepción de profundidad, reportó las siguientes observaciones sobre el área de la lámina esclerótica (mm^2) de cabezas de nervios ópticos humanos:
2.75 2.62 2.74 3.85 2.34 2.74 3.93 4.21 3.88
4.33 3.46 4.52 2.43 3.65 2.78 3.56 3.01

    a) Calcule $\sum x_i$ y $\sum x_i^2$

b) Utilice los valores calculados en el inciso (a) para determinar la varianza muestral S^2 y la desviación estándar muestral S.

c) Determine los cuartos inferior y superior.

d) Calcule el valor de la cuarta dispersión.

e) Si los dos valores muestrales más grandes, 4.33 y 4.52 hubieran sido 5.33 y 5.52, ¿Cómo afectará esto a f_s?. Explique.

f) Si a la muestra se agrega una décimo octava observación $x_{18} = 4.60$, ¿Cuánto vale f_s?.

Capítulo 2

Probabilidad

2.1 Ejercicios Resueltos

PROBLEMA 1[1]

La biblioteca de una universidad tiene cinco ejemplares de un cierto texto de reserva. Dos ejemplares(1 y 2) son primeras impresiones y las otras tres(3, 4 y 5) son segundas impresiones. Un estudiante examina estos libros en orden aleatorio, deteniéndose sólo cuando selecciona una segunda impresión.
Dos posibles resultados son 5 y 2, 1, 3.

a) Haga una lista de los resultados posibles.

b) Si A simboliza el evento cuando exactamente un libro es examinado, ¿cuáles resultados están en A?

c) Si B es el evento cuando el libro 5 es seleccionado, ¿cuáles resultados están en B?

d) Si C es el evento cuando el libro 1 no se examina, ¿cuáles resultados están en C?

SOLUCIÓN

a) $S = \{3, 4, 5, 13, 14, 15, 23, 24, 25, 123, 124, 125, 213, 214, 215\}$

b) $A = \{3, 4, 5\}$

c) $B = \{5, 15, 25, 125, 215\}$

d) $C = \{3, 4, 5, 23, 24, 25\}$

[1] TAV 2004

PROBLEMA 2[2]

Un sistema puede tener tres tipos de defectos: $A_i (i = 1, 2, 3)$ es cuando este sistema tiene un defecto del tipo i. Suponga que

$$\begin{array}{rclrclrcl} P(A_1) & = & 0.12 & P(A_2) & = & 0.07 & P(A_3) & = & 0.05 \\ P(A_1 \cup A_2) & = & 0.13 & P(A_1 \cup A_3) & = & 0.14 & P(A_2 \cup A_3) & = & 0.10 \\ P(A_1 \cap A_2 \cap A_3) & = & 0.01 & & & & & & \end{array}$$

a) ¿Cuál es la probabilidad de que el sistema no tenga el defecto tipo 1?

b) ¿Cuál es la probabilidad de que el sistema tenga defectos tipo 1 y 2 al mismo tiempo?

c) ¿Cuál es la probabilidad de que el sistema tenga defectos tipo 1 y 2 al mismo tiempo, pero que no tenga defectos tipo 3?

d) ¿Cuál es la probabilidad de que el sistema tenga dos de estos defectos?

SOLUCIÓN

a)

$$\begin{aligned} P(A_1^c) &= 1 - P(A_1) \\ &= 1 - 0.12 \\ &= 0.88 \end{aligned}$$

b)

$$\begin{aligned} P(A_1 \cap A_2) &= P(A_1) + P(A_2) - P(A_1 \cup A_2) \\ &= 0.12 + 0.07 - 0.13 \\ &= 0.06 \end{aligned}$$

c)

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3^c) &= P(A_1 \cap A_2) - P(A_1 \cap A_2 \cap A_3) \\ &= 0.06 - 0.01 \\ &= 0.05 \end{aligned}$$

[2]TAV 2004

d)

$$\underbrace{1 - P(A_1 \cap A_2 \cap A_3)}_{\text{prob. de a lo más dos errores}} = 1 - 0.01 = 0.99$$

PROBLEMA 3[3]

Cierto télefono público(que usalmente falla) devuelve la moneda insertada con probabilidad 0.6; hace la conexión con el número que uno marca con probabilidad 0.2 ; se queda con la moneda y no da la conexión requerida con probabilidad 0.3. Encuentre la probabilidad que una persona haga la llamada gratis.

SOLUCIÓN

A :"el teléfono público devuelve la moneda insertada".

B :"el teléfono público hace la conexión con el número que uno marca".

Se tiene

$P(A) = 0.6$, $P(B) = 0.2$, $P(A^c \cap B^c) = 0.3$

Se pide encontrar $P(B \cap A)$

$$\begin{aligned} P(B \cap A) &= P(B) + P(A) - P(A \cup B) \\ &= P(B) + P(A) - P[(A^c \cap B^c)^c] \\ &= P(B) + P(A) - \{1 - P[(A^c \cap B^c)]\} \\ &= 0.2 + 0.6 - \{1 - 0.3\} \\ &= 0.1 \end{aligned}$$

Luego la probabilidad que una persona haga la llamada gratis es 0.1.

PROBLEMA 4[4]

Se sabe que A contiene a B y que es disjunto con C. También se sabe que A es dos veces más probable que B, tres veces más probable que C y un medio veces más probable que su complemento, A^c. Encuentre $P(B \cup C)$ y $P(B \cap C)$.

[3] I1 segundo semestre 2003

[4] I1 y examen, segundo semestre 2003

SOLUCIÓN

Se tiene $B \subset A$, $A \cap C = \phi$, $P(A) = 2P(B)$

$P(A) = 3P(C)$, $P(A) = \frac{1}{2}\ P(A^c)$

De aquí obtenemos $P(A) = \frac{1}{3}$, $P(B) = \frac{1}{6}$, $P(C) = \frac{1}{9}$

Debemos calcular $P(B \cup C)$

$$\begin{aligned} P(B \cup C) &= P(B) + P(C) - P(B \cap C) \\ &= \frac{1}{6} + \frac{1}{9} - 0 \\ &= \frac{5}{18} \end{aligned}$$

$P(B \cap C) = 0$, ya que $B \subset A$ por lo tanto es disjunto con C.

PROBLEMA 5[5]

Una muestra aleatoria de 200 mujeres que toman anticonceptivo oral y se controlan en el Servicio Obstétrico y Ginecológico de cierto hospital fueron clasificadas según grupo sanguíneo (GS) y si presentan o no tromboembolismo (T) según se muestra en la tabla siguiente:

Grupo Sanguíneo	Número de Tromboelismo	Mujeres Sanas
A	32	51
B	8	19
AB	6	5
O	9	70
Total	55	145

Si selecciona en forma aleatoria a una mujer de esta muestra. Calcule la probabilidad de que:

a) Su GS sea A.

b) Su GS sea O y se presente Tromboembolismo.

c) Su GS sea B dado que la mujer está sana.

d) Su GS no sea A y presente Tromboebolismo.

SOLUCIÓN

Completando la tabla con los totales se tienen:

[5] segundo semestre de 2000

Grupo Sanguíneo(GS)	Número de Tromboelismo (T)	Mujeres Sanas (MS)	Total
A	32	51	83
B	8	19	27
AB	6	5	11
O	9	70	79
Total	55	145	200

a) $P(GS_A) = \frac{83}{200} = 0.415$

b) $P(GS_O \cap T) = \frac{9}{200} = 0.045$

c)

$$
\begin{aligned}
P(GS_B|MS) &= \frac{P(GS_B \cap MS)}{P(MS)} \\
&= \frac{19/200}{145/200} \\
&= 0.1310
\end{aligned}
$$

d)

$$
\begin{aligned}
P(GS_A^2 \cap T) &= \frac{8+6+9}{200} \\
&= 0.115
\end{aligned}
$$

PROBLEMA 6[6]

Suponga que cierto rasgo oftálmico está asociado con el color de ojos. Se estudiaron 300 individuos elegidos aleatoriamente, con los resultados siguientes:

	Color de ojos			
Rasgos	Azul	Café	Otro	Totales
Presencia	70	30	20	120
Ausencia	20	110	50	180
Totales	90	140	70	300

Calcular la probabilidad de que una persona elegida aleatoriamente:

a) Tenga los ojos azules.

b) Tenga el rasgo.

[6] I2 recuperativa, segundo semestre de 2000

c) No tenga el rasgo y presente ojos café.

d) Tenga el rasgo dado que presentó otro color de ojos.

SOLUCIÓN

Definamos los siguientes eventos

A	:	Tenga ojos azules	C	:	Tenga ojos café
R	:	Tenga el rasgo	O	:	Tenga otro color de ojos

a) $P(A) = \frac{90}{300} = 0,3$

b) $P(R) = \frac{120}{300} = 0,4$

c) $P(R^c \cap C) = \frac{110}{300} = 0,366$

d) $P(R|O) = \frac{P(R \cap O)}{P(O)} = \frac{20/300}{70/300} = 0.2857$

PROBLEMA 7[7]

En una prueba de Métodos Estadísticos un alumno dispone de siete lápices para realizar ésta prueba, de los cuales dos son de color rojo y cinco de color azul. El alumno selecciona un lápiz al azar y enseguida extrae el otro de los restantes.

(a) Defina los eventos involucrados

(b) ¿Cuál es la probabilidad que el primer lápiz extraído sea azul y el segundo sea de color rojo?

(c) ¿Cuál es la probabilidad que el segundo lápiz extraído sea rojo?

SOLUCIÓN

a) A:"El primer lápiz extraído es de color azul".
B:"El segundo lápiz extraído es de color rojo".

b) Se pide calcular $P(A \cap B)$ y se tiene $P(A) = \frac{5}{7}$ y $P(B|A) = \frac{2}{6}$, luego

$$\begin{aligned} P(A \cap B) &= P(A) \times P(B|A) \\ &= \frac{5}{7} \times \frac{2}{6} \\ &= \frac{5}{21} \end{aligned}$$

[7] I1 segundo semestre de 2003

c) Aquí se pide calcular $P(B)$, en donde $B = (A \cap B) \cup (A^c \cap B)$ y, como $A \cap B, A^c \cap B$ son disjuntos, se tiene:

$$\begin{aligned} P(B) &= P(A \cap B) + P(A^c \cap B) \\ &= \frac{5}{21} + P(A^c) \times P(B|A^c) \\ &= \frac{5}{21} + \frac{2}{7} \times \frac{1}{6} \\ &= \frac{6}{21} \end{aligned}$$

PROBLEMA 8[8]

En una industria de productos Químicos, las unidades son producidas por tres líneas en proporciones 25:35:40. Un 5% un 4% y un 2% de las unidades producidas por cada línea, respectivamente, son defectuosos. Las unidades son mezcladas y enviadas a los compradores.

a) Determine la probabilidad que una unidad escogida al azar sea defectuosa.

b) Si un cliente encuentra una unidad defectuosa, determine la probabilidad que se haya producido en la primera línea.

SOLUCIÓN

a) Los eventos son:

D :“ unidad defectuosa ”
L_i : “ línea de producción i ”; $i = 1, 2, 3$

$$\begin{aligned} P(D) &= \sum_{j=1}^{3} P(D|L_j)P(L_j) \\ &= 0.05 \times 0.25 + 0.04 \times 0.35 + 0.02 \times 0.40 \\ &= 0.0345 \end{aligned}$$

(b) La probabilidad que se haya producido en la primera línea, dado que la unidad es defectuosa.

$$\begin{aligned} P(L_1|D) &= \frac{P(D|L_1)P(L_1)}{P(D)} \\ &= \frac{0.05 \times 0.25}{0.0345} \\ &= 0.36 \end{aligned}$$

[8] I1 segundo semestre 2003

PROBLEMA 9[9]

Un paciente de cáncer está siendo tratado con una combinación de tres fármacos. Se observa que cuando se utilizan simultáneamente, la probabilidad de que dos de los tres fármacos se inactiven es de 1/3, resultando que sólo uno de ellos permanece activo frente al tumor. La efectividad de cada fármaco, con respecto a producir una remisión del tumor, es diferente. El fármaco A se ha mostrado efectivo en un 50% de los casos; el fármaco B, en un 75%, y el fármaco C, en un 60%.
La enfermedad remite en el paciente. ¿Cuál es la probabilidad de que el responsable de ello sea el fármaco B?

SOLUCIÓN

Se tiene los eventos

R :Remisión del tumor
A :Fármaco A
B :Fármaco B
C :Fármaco C

$P(A) = P(B) = P(C)(= 1/3)$, $P(R|A) = 0.5$, $P(R|B) = 0.75$, $P(R|C) = 0.6$

$$P(B|R) = \frac{P(R|B)P(B)}{P(R|A)P(A) + P(R|B)P(B) + P(R|C)P(C)}$$

$$= \frac{0.75}{0.5 + 0.75 + 0.6}$$

$$= \frac{0.75}{1.85}$$

$$= 0.4054$$

PROBLEMA 10[10]

Sólo 1 de 1000 adultos está afectado por una rara enfermedad, para la cual se ha desarrollado una prueba de diagnóstico. Durante la prueba, cuando un individuo padece la enfermedad presentará un resultado positivo 99% de las veces, mientras que un individuo sin la enfermedad mostrará un resultado de prueba positivo sólo en 2% de las veces. Si se hace una prueba a un individuo seleccionado al azar y el resultado es positivo,¿cuál es la probabilidad de que el individuo tenga la enfermedad?

[9] TAV 2004
[10] TAV 2004

SOLUCIÓN

Sea

IE	:	individuo enfermo
IS	:	individuo sano
RP	:	resultado positivo de la prueba

Del enunciado se tiene:

$$\begin{array}{rclrcl} P(IE) & = & 0.001 & P(IS) & = & 0.999 \\ P(RP|IE) & = & 0.99 & P(RP|IS) & = & 0.02 \end{array}$$

Se pide $P(IE|RP)$

$$\begin{aligned} P(IE|RP) &= \frac{P(IE \cap RP)}{P(RP)} \\ &= \frac{P(RP|IE)P(IE)}{P(RP)} \\ &= \frac{0.99 \times 0.001}{0.02097} \\ &= 0.047 \end{aligned}$$

PROBLEMA 11

Se dispone de tres dados A, B y C. El dado A es equilibrado, mientras que B está cargado a favor de los números impares y C lo está a favor de los pares. Sea $p > \frac{1}{2}$(resp. $q < \frac{1}{2}$) la probabilidad de obtener un número impar al lanzar el dado B(resp. el dado C). El experimento consiste en elegir uno de los dados de acuerdo al mecanismo que se indica a continuación y luego lanzarlo tres veces(*simpre el mismo dado*). El mecanismo de selección consiste en lanzar una moneda no equilibrada D(con probabilidad de cara α) y seleccionar el dado A si sale cara; de salir sello se elige B o C con igual probabilidad.

a) Calcule la probabilidad de que en el primer lanzamiento del dado aparezca un número par.

b) Calcule la probabilidad de que el dado A haya sido seleccionado en la primera etapa si los dos primeros números obtenidos son impares.

c) Calcule la probabilidad de obtener un número impar en el tercer lanzamiento si se ha obtenido un número impar en los dos anteriores.

SOLUCIÓN

a) Sea

A: dado equilibrado
B: dado cargado $P(impar) = p > 1/2$
C: dado cargado $P(impar) = q < 1/2$

Moneda D $\rightarrowtail P(cara) = \alpha$, $cara \longrightarrow A$, $sello \longrightarrow B$ ó C

Primera forma

Sean

H: La moneda D mostró cara
I_i: En el lanzamiento $i-$ésimo aparece un número par

$$
\begin{aligned}
P(I_1) &= P(I_1|H)P(H) + P(I_1|H^c)P(H^c) \\
&= \frac{1}{2} \times \alpha + [P(I_1|B)P(B) + P(I_1|C)P(C)] \times (1-\alpha) \\
&= \frac{\alpha}{2} + \left[(1-p) \times \frac{1}{2} + (1-q) \times \frac{1}{2}\right] \times (1-\alpha) \\
&= 1 - \left(\frac{\alpha}{2} + \left(\frac{1-\alpha}{2}\right)(p+q)\right)
\end{aligned}
$$

Segunda forma

Sean p_A, p_B y p_C las probabilidades de elegir A, B y C respectivamente $p_A = \alpha$, $p_B = p_C = \frac{1-\alpha}{2} = \beta$

$$
\begin{aligned}
P(I_1) &= 1 - P(I_1^c) \\
&= 1 - \left(\frac{\alpha}{2} + \beta \times p + \beta \times q\right) \\
&= 1 - \left(\frac{\alpha}{2} + \beta \times (p+q)\right)
\end{aligned}
$$

b) Sea A:el dado dado A es seleccionado

Primera forma: Usando el teorema de Bayes

$$P(A|I_1^c \cap I_2^c) = \frac{P(I_1^c \cap I_2^c|A)P(A)}{P(I_1^c \cap I_2^c|A)P(A) + P(I_1^c \cap I_2^c|B)P(B) + P(I_1^c \cap I_2^c|C)P(C)}$$

$$= \frac{\frac{1}{4} \times \alpha}{\frac{1}{4} \times \alpha + p^2 \times \frac{1-\alpha}{2} + q^2 \times \frac{1-\alpha}{2}}$$

$$= \frac{1}{1 + \frac{2(1-\alpha)}{\alpha}(p^2 + q^2)}$$

Segunda forma

$$P(A|I_1^c \cap I_2^c) = \frac{P(A \cap I_1^c \cap I_2^c)}{P(I_1^c \cap I_2^c})$$

$$= \frac{\alpha \times \frac{1}{2} \times \frac{1}{2}}{\alpha \times \frac{1}{2} \times \frac{1}{2} + \beta \times p^2 + \beta \times q^2}$$

$$= \frac{1}{1 + \frac{4\beta}{\alpha}(p^2 + q^2)}$$

c)

$$P(I_3^c|I_1^c \cap I_2^c) = \frac{P(I_3^c \cap I_1^c \cap I_2^c)}{P(I_1^c \cap I_2^c)}$$

$$= \frac{\alpha \times \left(\frac{1}{2}\right)^3 + \frac{1-\alpha}{2} \times p^3 + \frac{1-\alpha}{2} \times q^3}{\frac{\alpha}{4} + \frac{1-\alpha}{2}(p^2 + q^2)}$$

2.2 Ejercicios Propuestos

2.2.1 Eventos y formas de contar

1. Todos los días, un niño dispone de 30 diarios para vender en la misma esquina. Defina un espacio muestral para el experimento, que consisite del número de ventas en un día cualquiera. Defina además los eventos:

   A: vende al menos cinco diarios.

   B: vende exactamente cinco diarios.

   C: vende a lo más cinco diarios.

2. Considerando el ejercicio anterior y si ahora, el experimento consisite en registrar el número de ventas que el hace en dos días sucesivos. Defina un espacio muestral razonable para este experimento y describa los eventos:

   A: vende al menos cinco diarios el primer día.

   B: vende al menos cinco diarios el segundo día.

   C: vende al menos cinco diarios ambos dias.

3. Considere el juego del lanzamiento de dos dados ordinarios.

   a) Determine el espacio muestral asociado.

   b) ¿Cuántos eventos puede usted definir?

   c) Describa los siguientes eventos:

      A: la suma de los dos dados es menor o igual a 3.

      B: el segundo dado muestra el número 6.

      C: el segundo dado muestra un número par.

4. Un animal muere(M) o sobrevive(S) en el curso de un experimento quirúrgico. El experimento será ejecutado con dos animales, si ambos sobreviven, no se realizará ningún otro ensayo, si exactamente un animal sobrevive, a uno más se le aplicará el experimento. Finalmente si ambos animales mueren, serán tratados dos adicionales. Determine el espacio muestral asociado al experimento.

5. Considere el experimento aleatorio siguiente: Una moneda es lanzada hasta obtener cara por primera vez.

   a) Describa el espacio muestral asociado a este experimento.

   b) Describa los siguientes eventos:

      A: la primera cara ocurre en tres o menos lanzamientos.

      B: un número impar de lanzamientos es necesario para obtener cara por primera vez.

6. Se lanza una moneda hasta que aparezca un total de dos sellos, pero no se hacen más de 4 lanzamientos. Liste todos los resultados de un espacio muestral apropiado para este experimento.

7. Una urna contiene 5 fichas, marcadas del 1 al 5. Se extraen tres fichas, liste todos los resultados que constituyen el evento "el segundo número más grande extraído fue 3".

8. En el juego del "crapito", un jugador lanza un par de dados. Si obtiene una suma de 7 u 11, él gana inmediatamente; si él obtiene una suma de 2, 3 ó 12 él pierde inmediatamente. De otra manera, si la suma es `x`, él continua lanzando hasta obtener un mismo puntaje `x`(en cuya caso gana)u obtener un 7(en cuyo caso pierde). Describa el evento "el jugador gana con un puntaje de 5".

9. Determine el número de elementos en los conjuntos i)-iii) y haga una lista de ellos.

   i.- Un comité de dos personas se va a formar de un grupo de 5 estudiantes: Patricia, Ricardo, Eduardo, Romina y Joaquín.

   ii.- Se va elegir un presidente y tesorero del grupo de estudiantes de arriba. El presidente debe ser mujer.

   iii.- Los cinco estudiantes arriba van a bailar. ¿Cuántas parejas(varón y dama)se pueden formar?

10. Un experimentador investiga el efecto de tres variables: presión, temperatura y el tipo de catalítico sobre el rendimiento en un proceso de refinado. Si el experimentador intenta usar tres niveles para la temperatura, tres niveles para la presión y dos tipos de catalítico, ¿cuántos ensayos experimentales tendrá que realizar si quiere considerar todas las combinaciones de presión, temperatura y tipos de catalíticos?

11. ¿Cuántos números se pueden formar al arreglar los dígitos del número 4130131(excluyendo los que comienzan por 0?

12. Una Galería de Arte tiene que montar una exposición, para ello tiene 10 cuadrados distintos;

    a) Si tiene uno sólo para 8 cuadros. ¿De cuántas formas distintas puede organizar la Exposición?
    Resp: 1814400

    b) Si tiene otro sólo con 6 espacios. ¿De cuántas formas puede organizar la Exposición?
    Resp: 151200

13. Un curso de Métodos Estadístico de 40 alumnos tiene que formar una comisión para organizar un congreso;

    a) Si la comisión debe ser de 4 alumnos. ¿Cuántos comités distintos se podrían formar?
    Resp: 91390

b) Si la comisión debe ser de 8 alumnos. ¿Cuántos comités distintos se podrían formar?
Resp: 769046685

14. En una baraja de naipes se reparten 4 naipes a cada jugador:

    a) ¿Cuántas manos distintas pueden formarse?
    Resp: 316251

    b) ¿Cuántas manos de 20 Ases y 2 Queinas?
    Resp: 36

    c) ¿Cuántas manos distintas de 3 Reyes y 1 Jota?
    Resp: 16

    d) ¿Cuántas manos distintas con 2 Jotas y 2 cartas cualquiera?
    Resp:7956

15. Un testigo de un accidente de tránsito en el que el causante huyó, le indica a un policía que el número de matrícula del automóvil tenía las letras RLH seguidas por tres dígitos, el primero de los cuales era un cinco. Si el testigo no puede recordar los otros dos dígitos pero está seguro de que los tres eran diferentes, encuentre el número máximo de registros de automóvil que debe verificar la policia.
Resp: 72

16. Cuatro matrimonios compraron ocho lugares para un concierto. ¿En cuántas formas diferentes pueden sentarse?:

    a) Sin restricciones
    Resp: 40320

    b) Si se sientan por parejas
    Resp: 384

    c) Si todos los hombres se sientan juntos a la derecha de todas las mujeres
    Resp: 576

17. ¿Dé cuántas maneras puede un cliente ordenar un emparedado y una bebida si hay cinco clases de emparedado y una bebida si hay cinco clases de emparedado y cuatro de bebidas en el menú?
Resp:20 maneras

18. A diez ratas de laboratorio se le asignan rangos por su habilidad para aprender cinco diferentes tareas. A una rata se le asigna los valores tres, dos, uno y cero por cada tarea si ella es aprendida en uno, dos, tres o más períodos de entrenamiento.

    a) ¿Cuántos valores totales son posibles para una rata en particular?

    b) Suponiendo que no hay empates(dos ratas no puden recibir el mismo puntaje total)¿cuántos rangos son posibles? Resp:a) 16 valores b)10! rangos

19. Una empresa alimenticia desea numerar las fichas de sus clientes con números de cuatro cifras. Para ello dispone de los dígitos del cero al nueve. ¿Cuántas fichas podrá numerar si

    a) los dígitos pueden repetirse?
    b) los dígitos no pueden repetirse?
    c) el último número debe ser cero y los dígitos no pueden repetirse?
    Resp: a)1000 fichas b)5040 fichas c)504 fichas.

20. a) Determine el número de palabras de cuatro letras que se pueden formar con las letras de la palabra HEMBRAS.
    b) ¿Cuántas de ellas contienen sólo consonantes?
    c) ¿Cuántas empiezan y terminan con consonante?
    d) ¿Cuántas empiezan con vocal?
    e) ¿Cuántas contienen la letra B?
    f) ¿Cuántas empiezan con H y terminan con vocal?
    g) ¿Cuántas contienen ambas vocales?
    Resp: a)840 palabras b) 120 palabras c) 400 palabras d)240 palabras e) 480 palabras f)40 palabras g)240 palabras.

21. ¿Cuántos códigos de cuatro letras se pueden obtener con las primeras diez letras del alfabeto si

    a) ninguna letra se puede repetir?
    b) se pueden repetir las letras las veces que se desee?
    c) en b) las letras abyacentes no pueden ser iguales?
    Resp:a)5040 códigos b)10000 códigos c)7290 códigos.

22. Considere los siguientes dígitos: uno, dos, tres, cinco y seis

    a) ¿Cuántos números de tres cifras, se pueden formar si no se permiten repatición?
    b) ¿Cuántos de los obtenidos en a) son menores de 500?
    Resp:a)60 números b)36 números.

23. En la formación de un rasgo de un individuo intervienen cinco tipos de genes distintos. ¿Cuántos genotipos distintos se pueden obtener?.
    Resp:15 genotipos.

24. Suponga que hay seis síntomas reconocidos para una cierta enfermedad. La enfermedad puede ser diagnosticada si el paciente muestra cuatro ó más de los síntomas. ¿Dé cuántas maneras pueden combinarse los síntomas para poder dar un diagnóstico?
    Resp: 22 maneras

25. ¿Cuántas señales de tres luces distintas pueden obtenerse con cinco luces distintas?
    Resp:60 señales

26. Un médico visita seis salas diferentes de un hospital todos los días. Con el objeto de impedir a los pacientes que sepan cuando los visitará, varía el orden de las visitas. ¿De cuántas maneras distintas puede realizar sus visitas el médico?
Resp:720 maneras.

27. a) ¿Cuántas palabras diferentes se pueden formar con todas las letras de la palabra ESTADISTICA?
    b) ¿Cuántas de ellas empiezan y terminan con S?
    Resp:a)249480 palabras b)45360 palabras.

28. En un laboratorio se dispone de cinco ratas blancas y ocho grises. En un experimento deben participar cuatro ratas de las cuales una debe ser blanca. ¿Cuántas maneras distintas tiene el investigador de elegir las ratas?
Resp:280 maneras.

29. Se dispone de doce claveles de colores distintos y cinco rosas de colores distintos. Se desea formar un ramo de seis flores. ¿Cuántos ramos distintos se pueden formar si se desea que el ramo contenga al menos tres rosas?
Resp:11110 ramos.

30. Un grupo de profesionales está formado por seis médicos, tres farmacéuticos y cuatro biólogos. Se nombra una comisión de cinco personas para asistir a un congreso.
    a) ¿Cuántas comisiones distintas se pueden formar?
    b) ¿Cuántas comisiones distintas se pueden formar
        i) si el presidente y el secretario del grupo van por derecho propio?
        ii) si asiste por lo menos un farmacéutico?
    c) si asiste dos médicos, dos biólogos y un farmacéutico?
    Resp:a)1287 comisiones b) i) 165 comisiones ii)1035 comisiones c)270 comisiones.

31. Un estudiante debe contestar ocho de diez preguntas de un examen.
    a) ¿Cuántas maneras de escoger tiene?
    b) ¿Cuántas si las tres primeras preguntas son obligatoria?
    c) ¿Cuántas si tiene que contestar por lo menos cuatro de las cinco primeras preguntas?
    Resp:a)45 maneras b)21 maneras c)35 maneras.

32. Un estudiante tiene once amigos.
    a) ¿Dé cuántas maneras puede invitar a cinco de ellos?
    b) ¿Dé cuántas maneras, si dos de ellos están enojados y no asisten juntos?
    Resp:a) 462 maneras b)378 maneras.

33. De un grupo de ocho farmacéuticos, tres biólogos y cuatro médicos se elige una comisión de cinco personas. ¿Cuántas comisiones

a) distintas se pueden formar?

b) tiene exactamente dos médicos?

c) tiene por lo menos tres farmacéuticos?

d) tiene a un farmacéutico determinado?

e) tiene a lo más un biólogo?
Resp: a)3003 comisiones b)990 comisiones c) 1722 comisiones d)1001 comisiones e)2277 comisiones.

34. Un sicólogo está tratando a catorce pacientes, seis hombres y ocho mujeres. Debe seleccionar nueve pacientes para un experimento en grupo.

a) ¿Cuántos grupos puede formar el sicólogo?

b) ¿Cuántos grupos puede formar que

i) contengan cuatro hombres?

ii) incluyan a lo más dos mujeres?
Resp:a)2002 grupos b)i)840 grupos ii)0 grupo.

2.2.2 Probabilidad Condicional, Independencia y Teorema de Bayes

1. [11] Los eventos A y B son disjuntos y uno de ellos debe ocurrir. Sus probabilidades son p y p^2 respectivamente. Encuentre p.
   Resp:$p = (-1 + \sqrt{5})/2$

2. Suponga que$P(A) = 0.4$, $P(B^c) = 0.6$, $P(A \cap B^c) = 0.1$. Encuentre $P(A \cup B^c)$.

3. Se sabe que A contiene a B y que es disjunto con C. También se sabe que A es dos veces más probable que B, tres veces más probable que C y un medio más probable que su complemento, A^c. Encuentre $P(B \cup C)$ y $P(B \cap C)$.
   Resp:$P(B \cup C) = 5/18$ y $P(B \cap C) = 0$

4. Un grupo de alumnos siguió tres asignaturas: A, By C. Se sabe que: El 5% aprobó las tres asignaturas, el 45% aprobó A ó B, el 25% aprobó A y C, el 80% no aprobó B, el 40% aprobó A y el 35% aprobó C. Si se elige al azar un alumno de éste grupo.
   Calcule la probabilidad de que:

   a) Haya aprobado sólo A.
   b) Haya aprobado todas las asignaturas, sabiendo que aprobó A y B.
   c) Haya reprobado B, sabiendo que también reprobó A.
      Resp:a)0.5 b) 0.3 c)0.167

5. Sean A, B y C tres sucesos de un espacio probabilístico (Ω,A,P) tales que $P(A) = 0.4$, $P(C) = 0.3$, $P(A \cap B) = 0.1$, $P(A \cup B) = 0.5$, $P(A \cap B \cap C) = \phi$

   a) Calcular la probabilidad de que no ocurre ninguno.
   b) Determine la probabilidad de que ocurren A y B, pero no C.
      Resp: a)0.2 b)0.1

6. Un vendedor de autos nuevos ha comprobado que los clientes solicitan en especial algunos de los siguientes extras: transmición automática (A), neumáticos panteros(B) y radio(C). Si el 70% de los clientes solicitan A, el 75% solicitan B, el 80% solicitan C , el 80% requieren A o B, el 85% requieren A o C, el 90% requieren B o C y el 95% requieren A o B o C. Calcular la probabilidad que:

   a) El próximo cliente solicite las tres opciones.
   b) El próximo cliente solicite sólo una radio.
   c) El próximo cliente solicite sólo una de las tres opciones.
   d) El próximo cliente no solicite ningún extra especial.

[11] Examen recuperativo 2003

Resp: a) 0.25 b)0.8 c)0.3 d)0.05

Respuesta parte a)

Se sabe $P(A) = 0.4$, $P(B) = 0.65$, $P(C) = 0.8$, $P(A \cup B) = 0.8$, $P(A \cup C) = 0.85$, $P(B \cup C) = 0.9$, $P(A \cup B \cup C) = 0.95$

Nos interesa $P(A \cap B \cap C)$

$$\begin{aligned} P(A \cap B \cap C) =& P(A \cup B \cup C) - P(A) - P(B) - P(C) + P(A \cap B) + P(A \cap C) + P(B \cap C) \\ =& 0.95 - 0.4 - 0.65 - 0.8 + P(A \cap B) + P(A \cap C) + P(B \cap C) \qquad (*) \end{aligned}$$

Primero calculemos

$$\begin{aligned} P(A \cap B) =& P(A) + P(B) - P(A \cup B) \\ =& 0.4 + 0.65 - 0.8 \\ =& 0.25 \end{aligned}$$

y

$$\begin{aligned} P(A \cap C) =& P(A) + P(C) - P(A \cup C) \\ =& 0.4 + 0.8 - 0.85 \\ =& 0.35 \end{aligned}$$

y

$$\begin{aligned} P(B \cap C) =& P(B) + P(C) - P(B \cup C) \\ =& 0.65 + 0.8 - 0.9 \\ =& 0.55 \end{aligned}$$

Volviendo a (*) y por lo anterior se tiene $P(A \cap B \cap C) = 0.25$

7. Sean A, B y C tres sucesos de un espacio probabilístico (Ω,A,P) tales que $P(A) = 5/10$, ,$P(B) = 7/10$, $P(C) = 3/5$, $P(A \cap B) = 2/5$, $P(A \cap C) = 2/5$, $P(B \cap C) = 1/5$, $P(A \cap B \cap C) = 1/10$

   a) Calcular la probabilidad de que ocurra al menos uno de los sucesos anteriores.

   b) ¿Cuál es la probabilidad de que no ocurra ninguno de los sucesos anteriores?

   c) Determinar la probabilidad de que ocurran exactamente dos de éstos sucesos.
   Resp:a)0.9 b)0.1 c)0.8

Respuesta a)

Se pide $P(A \cup B \cup C)$

$$\begin{aligned} P(A \cup B \cup C) =& P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) \\ =& 5/10 + 7/10 + 3/5 - 2/5 - 2/5 - 1/5 + 1/10 \\ =& 0.9 \end{aligned}$$

8. Una caja de fusibles contiene 20 unidades, de los cuales 5 son defectuosas. Si tres de estos fusibles son tomados al azar, en sucesión y sin remplazo.
   a) ¿Cuál es la probabilidad que los tres sean defectuosos?
   b) Si cada una de las dos primeras se extrajo un defectuoso, ¿cuál es la probabilidad que el tercero extraído sea bueno?
   c) Si los dos primeros estaban buenos, ¿cuál es la probabilidad que el tercero extraído sea defectuoso?
   d) ¿Cuál es la probabilidad que los dos primeros sean buenos y el tercero defectuoso?
   Resp:a)0.0087 b)0.83 c)0.277 d)0.15

9. Un caja contiene fichas blancas y negras cada una de las cuales tiene grabadas una letra A o C, la composición de la caja es:

	N	B	Total
A	5	3	8
C	1	2	3
Total	6	5	11

   Calcular la probabilidad de:
   a) Obtener una ficha negra.
   b) Que una ficha sea negra si se sabemos que tiene una A.
   c) Obtener una ficha que tenga una A.
   d) Obtener una ficha negra con una A.
   e) Obtener una ficha que tenga una A sabiendo que es negra.
   Resp: a)6/11 b)5/8 c)8/11 d)5/11 e)5/11

10. Suponga que una una caja con n_1 fichas blancas y n_2 fichas negras, se extraen al azar dos fichas sin reemplazo.
   a) ¿Cuál es la probabilidad de que las dos fichas extraídas sean blancas?
   b) ¿Cuál es la probabilidad de obtener una ficha negra en la primera extracción y una blanca en la segunda extracción?

   Resp: a) $(\frac{n_1}{n_1+n_2}) \times (\frac{n_1-1}{n_1+n_2-1})$ b)$(\frac{n_2}{n_1+n_2}) \times (\frac{n_1}{n_1+n_2-1})$

11. Un grupo de alumnos está compuesto de cuatro Químicos y cinco Físicos. Se eligen al azar dos de ellos. ¿Cuál es la probabilidad que:

    a) el primero sea físico y el segundo químico?

    b) el segundo es químico?

    c) los dos sean físicos?

    d) Si se eligen cuatro alumnos, uno a continuación del otro. ¿Cuál es la probabilidad que los primeros sean físicos y los dos últimos químicos?

    e) ¿Cuál es la probabilidad que sean alternados las carreras?
    Resp:a)5/18 b)4/9 c)5/18 d)5/63 e)10/63

12. La probabilidad que un estudiante estudie para un examen final es 0.20. Si estudia , la probabilidad de que apruebe el examen es 0.8 en tanto que si no estudia, la probabilidad es de sólo 0.50.

    a) ¿Cuál es la probabilidad que dicho estudiante apruebe su examen final?.

    b) Dado que aprobó su examen, ¿cuál es la probabilidad que él haya estudiado?
    Resp:a) 0.56 b)0.286

13. Se usa un interruptor para cortar un flujo cuando este alcanza un cierto nivel de profundidad en un estanque. La confiabilidad del interruptor(probabilidad que trabaje cuando debe) se supone de 0.9. Un segundo tipo de interruptor es puesto en paralelo y su confiabilidad es 0.7. Los interruptores trabajan en forma independiente.

    a) ¿Cuál es la confiabilidad de la combinación de los interruptores?

    b) ¿Cuál es la probabilidad, que solo trabaje el primer interruptor?

    c) ¿Cuál es la probabilidad que sólo uno de los interruptores trabaje?
    Resp:a)0.97 b)0.27 c)0.34

14. Dos máquinas de una planta elaboran el 10% y el 90% de la producción total de cierto artículo. La probabilidad de producir un artículo defectuoso con dichas máquinas es 0.01 y 0.05 respectivamente. ¿Cuál es la probabilidad que un artículo tomado al azar de la producción de un día haya sido producido con la primera máquina, sabiendo que es defectuoso?

15. Sean A y B dos eventos asociados a un espacio muestral Ω, tales que: $P(A) = 1/4$, $P(B|A) = 1/2$ y $P(A|B) = 1/4$.

    a) ¿Son A y B eventos mutuamente excluyentes?

    b) ¿Es $A \subset B$?

    c) ¿Son A y B eventos independientes?

    d) Determine $P(A^c|B^c)$

16. La probabilidad que un alumno de un curso determinado se titule en cinco años es 3/5. La probabilidad que una alumna de dicho curso tenga su título en cinco años más es 5/8. Calcular:

    a) Probabilidad de que ambos se titulen en cinco años más.

    b) Probabilidad de que al menos uno de ellos lo haga.

    c) Probabilidad de que el alumno no se titule y la alumna sí.

Capítulo 3

Variable Aleatoria Discreta

3.1 Ejercicios Resueltos

PROBLEMA 1[1]

Usted tiene cinco monedas en su bolsillo; dos de un peso, dos de cinco pesos y una de diez pesos, donde las monedas de igual valor son distinguibles. Tres monedas son extraídas al azar. Sea X la cantidad extraída(en pesos). Encuentre:

a) La función de distribución acumulada de probabilidades de X.

b) $P(X \leq 10|X \leq 15)$.

c) La función generadora de momentos y calcule mediante ésta la esperanza y la desviación estándar de X e interprete.

SOLUCIÓN

a)

X	7	11	12	16	20
$P(X=x)$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{4}{10}$	$\frac{1}{10}$

b) $$P(X \leq 10|X \leq 15) = \frac{P(X\leq 10, X\leq 15)}{P(X\leq 15)} = \frac{\frac{2}{10}}{\frac{5}{10}} = \frac{2}{5}$$

c) Sabemos de a) que el recorrido de X esta dado por $R_X = \{7, 11, 12, 16, 20\}$

$$\mathcal{M}_X(t) = E[e^{tx}] = \sum_{x\in R_X} e^{tx} P(X=x)$$

luego la función generadora esta dado por:

[1] I2 segundo semestre 2003

$$\mathcal{M}_X(t) = \tfrac{1}{10}(2e^{7t} + 2e^{11t} + e^{12t} + 4e^{16t} + e^{20t})$$

calculemos ahora el primer y el segundo momento

$$\begin{aligned}
\mathcal{M}'_X(t) &= \tfrac{1}{10}(14e^{7t} + 22e^{11t} + 12e^{12t} + 64e^{16t} + 20e^{20t}) \\
\mathcal{M}''_X(t) &= \tfrac{1}{10}(14 \times 7e^{7t} + 22 \times 11e^{11t} + 12 \times 12e^{12t} + 64 \times 16e^{16t} + 20 \times 20e^{20t})
\end{aligned}$$

evaluando ahora en cero obtenemos la esperanza y varianza

$$\begin{aligned}
E[X] &= \mathcal{M}'_X(0) \\
&= 13.2 \\
V[X] &= \mathcal{M}''_X(0) - \{\mathcal{M}'_X(0)\}^2 \\
&= 190.8 - (13.2)^2 \\
&= 16.56
\end{aligned}$$

PROBLEMA 2[2]

Considere la v.a. X con función de distribución acumulada $F(x)$ dada por:

X	-2	-1	0	1	2
$F(x)$	0.1	0.3	0.5	0.7	1

a) Determine la función de distribución de la variable alatoria X.

b) Defina $Y = X^2 + 1$

 i) Determine la función de distribución acumulada de la variable aleatoria asociada a Y.

 ii) Calcule $E[X]$ y $V[X]$.

SOLUCIÓN

a)

X	-2	-1	0	1	2
$P(X = x)$	0.1	0.2	0.2	0.2	0.3

b) i) $P(Y = 1) = P(X = 0) = 0.2$;
$P(Y = 2) = P(X = -1) + P(X = 1) = 0.2 + 0.2 = 0.4$;
$P(Y = 5) = P(X = 2) + P(X = -2) = 0.3 + 0.1 = 0.4$

De aquí se obtiene la función de probabilidad para Y

[2] I2 Segundo semestre 2001

Y	1	2	5
$P(Y=y)$	0.2	0.4	0.4

La función de distribución es:

Y	1	2	5
$F(y)$	0.2	0.6	1

ii) $E[Y] = \sum_y yp(y) = 1 \times 0.2 + 2 \times 0.4 + 5 \times 0.4 = 3$

$E[Y^2] = \sum_y y^2 p(y) = 1^2 \times 0.2 + 2^2 \times 0.4 + 5^2 \times 0.4 = 11.8$

$V[Y] = E[Y^2] - (E[Y])^2 = 11.8 - 3^2 = 2.8$

PROBLEMA 3[3]

Para promocionar sus helados de paleta, una fábrica pone cada 15 helados una etiqueta que dice "vale otro". Cualquiera persona que compre un helado y le salga "vale otro$^{\underline{o}}$btiene uno gratis. Estos helados cuestan 100 pesos cada uno. Si Ud. decide comprar estos helados hasta obtener uno gratis. ¿Cuánto esperaría gastar?

SOLUCIÓN

Sea X : número de helados hasta obtener uno gratis, $X \sim G(p = \frac{1}{15})$

$E[X] = \frac{1}{p} = 15$ helados

pero como cada helado cuesta \$100, se esperaría gastar \$1500.

PROBLEMA 4[4]

Un basquetbolista efectúa repetidos lanzamientos desde una línea de tiros libres. Suponga que sus lanzamientos son ensayos de Bernoulli independientes con probabilidad de acertar de 0.7.

a) ¿Cuál es la probabilidad que le tome menos de cinco lanzamientos para efectuar un segundo acierto?

b) ¿Cuál es la probabilidad que le tome menos de cinco lanzamientos para efectuar su primer acierto?

[3] I2 segundo semestre de 2003
[4] I2 segundo semestre de 2003

c) ¿Cuál es el número esperado de lanzamientos para lograr su cuarto acierto?

SOLUCIÓN

a) Y :"número de lanzamientos hasta el segundo acierto", $Y \sim Bineg(r = 2, p = 0.7)$

$$\sum_{y=2}^{4} \binom{y-1}{1} 0.7^2 0.3^{y-2} = 0.9163$$

b) X :"número de lanzamientos hasta el primer acierto", $X \sim G(p = 0.7)$

$$P(X < 5) = 1 - P(X \geq 5) = 1 - P(X = 5) = 1 - 0.3^4 = 0.9919$$

c) Z :"número de lanzamientos hasta el cuarto acierto", $Z \sim Bineg(r = 4, p = 0.7)$

$$E[Z] = \frac{4}{0.7} = 5.7143$$

PROBLEMA 5[5]

En cierta área rural, una extraña enfermedad está afectando a uno de cada 100 niños. Además se observa que en promedio, aparece un caso cada 30 días.

a) Se tiene la información que en el sector existen un total de 300 niños, determine la probabilidad que la extraña enfermedad afecte tan sólo a 2 de ellos. Hacerlo de dos formas diferentes.

b) Determine la probabilidad que en un período de 15 días se observe 2 casos como mínimo.

SOLUCIÓN

a) **Primera Forma**

X :"número de casos afectados en 30 días", $X \sim P(\lambda = np = 300 \times 0.01 = 3)$

$P(X = 2) = \frac{3^2 e^{-3}}{2!} = 0.22404$

Segunda Forma

$X \sim Bin(n = 300, p = 0.01)$

$P(X = 2) = \binom{300}{2} 0.01^2 0.99^{298} = 44850 \times 0.0001 \times 0.0500 = 0.2244$

[5] TAV 2004

b) $\lambda = 1/2 \left(\lambda = 1 \left[\frac{casos}{30dias}\right] = \frac{1}{2}\left[\frac{casos}{15dias}\right]\right)$

$$\therefore P(X \geq 2) = 1 - P(X < 2) = 1 - e^{-1/2}\left(\frac{\frac{1}{2}^0}{0!} + \frac{\frac{1}{2}^1}{0!}\right) = 0.0902$$

PROBLEMA 6[6]

En las moscas de la fruta, cuatro de cada 10^5 espermatozoide presentan una mutuación del color rojo de los ojos a blanco, o viceversa.

a) ¿Cuántas mutuaciones esperaría usted que se produjesen en 200000 espermatozoide? Justifique

b) ¿Cuál es la probabilidad de que se produzcan entre 6 y 10, ambas inclusive para el caso de 200000 espermatozoide?

(Nota: Si $X \sim Bin(n,p)$ con $n > 30$ y $p \approx 0$, entonces $X \sim P(\lambda)$ con $\lambda = n \times p$)

SOLUCIÓN

$P(\text{presenta mutuación}) = \frac{4}{10^5} = 0.00004$

a) $n = 200000$ y $p = 0.00004$

X : número de mutuaciones, $X \sim Bin(n,p)$.

Luego $E(X) = n \times p = 8$

b)

$$P(6 \leq X \leq 10) = P(X=6) + P(X=7) + P(X=8) + P(X=9) + P(X=10)$$

Nota:

$$\begin{aligned} X \sim Bin(n = 200000, p = 0.00004) &\Rightarrow X \sim P(\lambda = 8) \\ &\Rightarrow P(X=k) = \frac{8^k e^{-8}}{k!}; k = 0,1,2,\ldots \end{aligned}$$

Luego $P(6 \leq X \leq 10) = e^{-8}\left\{\frac{8^6}{6!} + \frac{8^7}{7!} + \frac{8^8}{8!} + \frac{8^9}{9!} + \frac{8^{10}}{10!}\right\} = 0.6246$

PROBLEMA 7[7]

Señale si es verdadero(V) o falso(F), justifique cuando sea falso dando a conocer el verdadero significado o valor númerico según corresponda:

[6] I2, segundo semestre de 2000
[7] I2 segundo semestre de 2001

a) Si $X \sim Geo(p)$, entonces X es el número de la repetición en la cual se obtiene éxiste por 3 vez.

b) Si $X \sim Bin(n, p)$, entonces X es el número de veces que ocurre éxito en una cantidad $n(n:$ indefinida).

c) La v.a. X : número de ocurrencias en el intervalo o región tiene una distribución de $Poisson$ con tasa λ.

d) Será cierto que:
$$E[g(x)] = \sum_x g(x)P(X = x)$$

e) En el modelo Hipergeométrico el tipo de muestreo es con restitución(esto es, se toman elementos uno a uno y cada vez que se saca uno de ellos se devuelve).

f) Si $X \sim Poisson(4)$, entonces $P(X \geq 6) = 0.1042$.

SOLUCIÓN

a) Falso; X: n° de la repetición en la cual se obtiene éxito por primera vez.

b) Falso; n es conocido.

c) Verdadero.

d) Verdadero.

e) Falso; es sin reposición.

f) Falso;
$$P(X \geq 6) = P(X = 6) + P(X = 7) + \ldots$$
donde $P(X = 6) = \frac{4^6 e^(-4)}{6!} = 0.104195$, y $P(X = 7) = \frac{4^7 e^(-4)}{7!} = 0.05954$ lo que implica $P(X = 6) + P(X = 7) = 0.1637 > 0.1042$

PROBLEMA 8[8]

Diez individuos, cada uno de ellos propenso a la tuberculosis, entran en contacto con un portador de la enfermedad. La probabilidad de que la persona se contagie del portador a un sujeto cualquiera es 0.10.

a) ¿Cuál es la probabilidad de que por lo menos dos personas contraigan la enfermedad?

b) ¿Cuántos se espera que contraigan la enfermedad?

[8] I2 segundo semestre de 2001

SOLUCIÓN

a) X:“n° de personas que contraen la enfermedad. ($X \sim (10; 0.1)$)

$$\begin{aligned} P(X \geq 2) =& 1 - P(X < 2) \\ =& 1 - [P(X = 0) + P(X = 1)] \\ =& 1 - \left[\binom{10}{0} 0.1^0 0.9^{10} + \binom{10}{1} 0.1^1 0.9^9\right] \\ =& 0.2639 \end{aligned}$$

b) $E[X] = n \times p = 10 \times 0.2639 \approx 3$

PROBLEMA 9[9]

Un comité de 3 integrantes se forma aleatoriamente seleccionado de entre 4 doctores y dos enfermeras. (se eligen uno a uno las personas y estas no se vuelve a considerar en el siguiente paso)

a) Escriba la función de probabilidad para la variable aleatoria X que representa el número de doctores en el comité.

b) Encuentre $P(X \leq 4\sigma)$.

SOLUCIÓN

a) X : n° de doctores en el comité ($X \in \{1, 2, 3\}$)

$$P(X = 1) = \frac{\binom{4}{1}\binom{2}{2}}{\binom{6}{3}} = 0.2 \qquad P(X = 2) = \frac{\binom{4}{2}\binom{2}{1}}{\binom{6}{3}} = 0.6$$

Luego

X	1	2	3
$P(X = x)$	0.2	0.6	0.2

b)

$$E[X] = \sum_x xp(x) = 1 \times 0.2 + 2 \times 0.6 + 3 \times 0.2 = 2$$

$$E[X^2] = \sum_x x^2 p(x) = 1^2 \times 0.2 + 2^2 \times 0.6 + 3^2 \times 0.2 = 4.4$$

[9] I2 segundo semestre de 2001

$$V[X] = E[X^2] - (E[X])^2 = 4.4 - 2^2 = 0.4$$

$\Rightarrow \sigma = 0.6324555$

Luego

$$\begin{aligned} P(X \leq 4\sigma) &= P(X \leq 2.5298) \\ &= P(X = 1) + P(X = 2) \\ &= 0.2 + 0.6 \\ &= 0.8 \end{aligned}$$

PROBLEMA 10[10]

El número de infracciones expedidas por un lector de parquímetro puede modelarse mediante un modelo de Poisson con una tasa de cinco infracciones por hora.

a) ¿Cuál es la probabilidad de que exactamente cuatro infracciones se expidan durante una hora en particular?

b) ¿Cuál es la probabilidad de que por lo menos cuatro se expidan durante una hora en particular?

c) ¿Cuántas infracciones se espera expedir durante un período de 45 minutos?

SOLUCIÓN

X : número de infracciones expedidas por un lector de parquímetro $X \sim Poisson(\lambda = 5)$

a) $P(X = 4) = \frac{e^{-5} \times 5^4}{4!}$

b)

$$\begin{aligned} P(X \geq 4) =& 1 - P(X < 4) \\ =& 1 - P(X \geq 3) \\ =& 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)] \\ =& 1 - 0.265 \\ =& 0.735 \end{aligned}$$

c) $\lambda = 5 \times \frac{45}{60} = 3.75$

[10] TAV 2004

3.2 Problemas Propuestos

1. [11]Una gran empresa química compra varios componentes de laboratorio cada año, cuya cantidad depende de la frecuencia de reparaciones en el año anterior. Suponga que el número de componentes de laboratorio, X, que se compran cada año tiene la siguiente distribución de probabilidad.

x	0	1	2	3
$p(x)$	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{2}{5}$	$\frac{1}{5}$

   a) Calcular $E[X]$, $E[X^2]$ y $Var[X]$.

   b) Si el costo del modelo que se desea adquirir permanece sin cambio en \$1200 durante un año y se ofrece un descuento de $50X^2$ en cualquier compra, ¿Cuánto dinero espera esta firma invertir en componentes de laboratorio para fin de año?

2. [12]Cierta empresa envía 40% de sus paquetes de correo nocturnos por servicio de correo expreso E_1. De estos paquetes, 2% llega después de la hora garantizada de entrega (señalar con L el evento de *entregado tarde*).

   a) Si se selecciona al azar un registro de envíos nocturnos de los archivos de la compañía, ¿Cuál es la probabilidad de que el paquete sea enviado vía E_1 y llegue tarde?

   Suponga ahora que el 50% de los paquetes nocturnos son enviados por servicio de correo expreso E_2 y el restante 10% son enviados por E_3. De los enviados por E_2 sólo el 1% llega tarde, mientras que 5% de los paquetes manejados por E_3 llega tarde.

   b) ¿Cuál es la probabilidad de que un paquete seleccionado al azar llegue tarde?

   c) Si un paquete seleccionado al azar llega a tiempo, ¿cuál es la probabilidad de que no haya sido enviado por E_1?

3. [13]En cierto servicio telefónico, la probabilidad de que una llamada sea contestada en menos de 30 segundos es 0.75. Suponga que las llamadas son independientes.

   a) Si una persona llama 10 veces, ¿cuál es la probabilidad de que exactamente nueve de las llamadas sean contestadas en un espacio de 30 segundos?

   b) Si una persona llama 20 veces, ¿cuál es la probabilidad de que al menos 16 de las llamadas sean contestadas en menos de 30 segundos?

   c) Si una persona llama 20 veces, ¿cuál es el número promedio de llamadas que serán contestadas en menos de 30 segundos?

[11]I1 segundo semestre de 2002
[12]I1 segundo semestre de 2002
[13]I1 segundo semestre de 2002

d) ¿Cuál es la probabilidad de tener que llamar cuatro veces para obtener la primera respuesta en menos de 30 segundos?

e) ¿Cuál es el número promedio de llamadas necesario para obtener dos respuestas en menos de 30 segundos?

4. Un Asociación de químicos comienza una campaña telefónica con el propósito de aumentar el número de socios. Con base en experiencia previa, se sabe que una de cada 20 personas que reciben la llamada se une al club. Si en un día 25 personas reciben la llamada telefónica.

   a) ¿Cuál es la probabilidad de que por lo menos dos de ellas se inscriben en el Asociación?

   b) ¿Cuál es el número esperado de personas?
   Resp. a)0.3576 b)aprox 1.

5. La probabilidad de que una persona tenga una mala reacción a la inyección de determinado suero es 0.001, determine la probabilidad de que 20 individuos,

   a) exactamente tres tengan una reacción mala.

   b) Más de dos individuos tengan una reacción mala.

   c) ¿Cuál es el número de individuos que se espera que tengan una mala reacción?
   Resp: a)0.000001120 b)1.125×10^{-6} c)0.02

6. La Colorado Power Company proporciona tarifas más bajas a los clientes que prefieran las horas de menos consumo, el 30% de sus clientes aprovecha estos ahorros. El departamento de servicios a clientes ha elegido a 12 clientes al azar para que participen en un grupo de interés para discutir a qué horas se produce el mayor consumo de energía. Al departamento de supervisión le preocupa que el grupo contenga una gran proporción de usuarios que prefieran la tarifa baja.

   a) ¿Cuál es la probabilidad de obtener menos de tres usarios de tarifa baja en el grupo de interés?

   b) ¿Cuál es la probabilidad de obtener más de cuatro usuarios de tarifa baja en el grupo de interés?

   c) ¿Cuál es la probabilidad de obtener menos de ocho clientes normales en el grupo de interés?

   d) calcule la media y la desviación estándar para los usuarios de tarifa baja en el grupo de interés.
   Resp: a)0.253 b)0.276 c)0.275 d)3.6 y 1.59.

7. El 60% de los residentes de la región metropolitana se registraron para votar. Si se elige al azar 10 personas con edad para votar, encuentre la probabilidad de obtener:

   a) diez electores registrados

   b) exactamente cinco electores registrados

   c) ningún elector registrado
   Resp: a)0.006 b)0.201 c)0

8. Dado que no todos los pasajeros de una aerolínea abordan el vuelo para el que han reservado un lugar, la aerolínea vende 125 boletos para un vuelo de 120 pasajeros. La probabilidad de que un pasajero no aborde el vuelo es 0.10, y el comportamiento de los pasajeros es independiente.

   a) ¿Cuál es la probabilidad de que todos los pasajeros aborden el vuelo?

   b) ¿Cuál es la probabilidad de que el vuelo parta vacío?
   Resp: a)0.996 b)0.989

9. La probabilidad de que una muestra de aire contenga una molécula rara es 0.01. Si se supone que las muestras son independientes con respecto a la presencia de la molécula rara, ¿cuál es la probabilidad de que sea necesario analizar exactamente 125 muestras antes de detectar una molécula rara?
   Resp: 0.0029

10. La probabilidad de un alineamiento óptico exitoso en el ensamblado de un producto de almacenamiento óptico de datos es 0.8. Suponga que los ensayos son independientes.

    a) ¿Cuál es la probabilidad de que el primer alineamiento exitoso requiera exactamente cuatro ensayos?

    b) ¿Cuál es la probabilidad de que el primer alineamiento exitoso requiera como máximo cuatro ensayos?

    c) ¿Cuál es la probabilidad de que el primer alineamiento exitoso requiera al menos cuatro ensayos?
    Resp. a)0.0064 b)0.9984 c)0.008

11. Supóngase que el costo de efectuar en experimento químco es \$1000. Si el experimento falla, se incurre en un costo adicional de \$300 debido a ciertos cambios que deben efectuarse antes de que se intente un nuevo experimento. Si la probabilidad de éxito en cualquiera de los ensayos es 0.2, si los ensayos aislados son independientes y si los experimentos continúan hasta que se obtiene el primer resultado exitoso, ¿cuál es el costo esperado del procedimiento completo?
    Resp: \$6200

12. En cierta región, la probabilidad de que ocurra una tormenta con truenos en un día cualquier durante dos meses de verano es igual a 0.1. Suponiendo independencia de un día con otro, ¿cuál es la probabilidad de que la primera tormenta con truenos del verano ocurra el día 3 del segundo mes?(considere 1 mes=31 días)
    Resp:0.003

13. Si la probabilidad de que cierto examen dé una reacción positiva igual a 0.4, ¿cuál es la probabilidad de que ocurran menos de cinco reacciones negativas antes de la primera

positiva?
Resp:0.92

14. Una aeronave de alto rendimiento contiene tres computadoras idénticas. Sólo una de ellas se utiliza para controlar la nave; las otras dos son reservas que se activan en caso de falla en el sistema primario. Durante una hora de operación, la probabilidad de falla en el computador primario(o en cualquiera de los sistemas de reserva que se encuentre activo) es 0.0005. Si se supone que cada hora representa un ensayo independiente,

    a) ¿Cuál es la probabilidad de que las tres computadoras fallen durante un vuelo de cinco horas?
    b) ¿Cuál es el tiempo promedio de las tres computadoras fallen?
    Resp:a)1.249×10^{-9} b)6000 horas.

15. La escala electrónica de un proceso de llenado automático detiene la línea de producción después de haber detectado tres paquetes con un peso menor de lo especificado. Suponga que la probabilidad de llenar un paquete con un peso menor es de 0.001 y que cada operación de llenado es independiente.

    a) ¿Cuál es el número promedio de operaciones de llenado antes de que detenga la línea de producción?
    b) ¿Cuál es la desviación estándar del número de operaciones de llenado antes que se detenga la línea de producció?
    Resp:a)3000 b)1731.18

16. Un auditor elige tres cuentas al azar de un grupo de 10 para examinarlas con todo cuidado. La compañía a la que se le hace la auditoría sabe que cuatro de las cuentas de impuestos tiene errores. ¿Cuál es la probabilidad de que las tres cuentas elegidas no tengan errores?
    Resp: 0.167

17. Un lote de 25 tubos de ensayo se someten a inspección. El procedimiento consiste en extraer 5 al azar, sin reemplazo y someterlos a prueba. Si menos de 2 tubos fallan, el lote es aceptado.De otra manera el lote rechazado.Suponga que el lote contiene 4 defectuosos.

    a) ¿Cuál es la probabilidad que el lote sea aceptado?
    b ¿Cuál es el número esperados de tubos defectuosos en la muestra?
    Resp: a)0.8335 b)1

18. Un lote de piezas contiene 100 de un proveedor local de tubería, y 200 de un proveedor del mismo material, pero de otro estado. Si se eligen cuatro piezas al azar y sin reemplazo,

    a) ¿Cuál es la probabilidad que todas provengan del proveedor local?
    b) ¿Cuál es la probabilidad de que 2 o mas piezas de la muestra sean del proveedor local?

c) ¿Cuál es la probabilidad de que al menos una pieza sea del proveedor local?
Resp: a)0.0119 b)0.408 c)0.196

19. Suponga que el número de fallas en un alambre delgado esta descrito por una distribución de Poisson con una media de 2.3 fallas por milímetros.

    a) Determine la probabilidad de tener exactamente dos fallas en un milímetro de alambre.
    b) Determine la probabilidad de tener 10 fallas en dos milímetros de alambre.
    c) Determine la probabilidad de tener al menos una falla en dos metros de alambre.
    Resp:a)0.2652 b)0.01175 c)0.989948

20. Después de una prueba de laboratorio muy rigurosa con cierto componente químico, el fabricante determina que en promedio, sólo fallarán dos componentes antes de tener 1000 horas de prueba. Un comprador observa que son cinco las que fallan antes de las 1000 horas. Si el número de componentes que fallan es una v.a. de Poisson, ¿exista suficiente evidencia para dudar de la conclusión del fabricante?
Resp:0.0361

21. Suponga que en un cruce transitado ocurren de manera aleatoria e independiente dos accidentes por semana.

    a) Determinar la probabilidad de que ocurra un accidente en la próxima semana.
    b) ¿Cuál es la probabilidad que en período de dos semanas ocurran al menos dos accidentes?
    Resp: a)0.2707 b)0.908

22. Una industria acerera fabrica una muestra reducida de alambre delgado de cobre para ser ofertado a los clientes preferenciales. El inspector de calidad en unos de los controles habituales determinó que el número de fallas está descrito por una distribucción de Poisson con una media de 2.3 fallas por milímetro. Determine la probabilidad de:

    a) Tener exactamente dos fallas en un milímetro de alambre.
    b) Tener 10 fallas en cinco milímetros de alambre.
    c) Tener al menos una falla en dos milímetro de alambre.
    Resp:a)0.265 b)0.113 c)0.9899

Capítulo 4

Variable Aleatoria Continua

4.1 Ejercicios Resueltos

PROBLEMA 1[1]

Un estudio realizado por investigadores, con pacientes con altos niveles de colesterol, demostró que una dieta de 12 días basada en alimentos sin colesterol junto con un poco de ejercicio moderado, desencadena una baja en tales niveles en la sangre. Demostró además que la pérdida representa un comportamiento de tipo normal con una media del 15% y una desviación estándar del 2,5%.

a) Para los investigadores, la disminución de los niveles de colesterol por sobre el 12.5% se considera como un objetivo logrado en su totalidad. Si la disminución se ubica entre el 10.1% y 12.5%, el objetivo se ha logrado medianamente, en cambio si la disminución es por debajo del 10.1%, el objetivo no se ha logrado. Suponga que se someterán a este tipo de dieta 5400 pacientes, determine la cantidad de personas que caerán en cada categoría.

b) Para un grupo de 5 pacientes que realiza la dieta, determine la probabilidad que 3 de ellos logren el objetivo en su totalidad

SOLUCIÓN

a) La figura 4.1 muestra la disminución de los niveles de colesterol

$n = 5400$ pacientes, X :"perdida en los niveles de colesterol", $X \sim N(15\%, (2.5\%)^2)$

Definamos L:logrado, NL: no logrado, ML: medianamente logrado

[1] I2 recuperativa segundo semestre de 2000

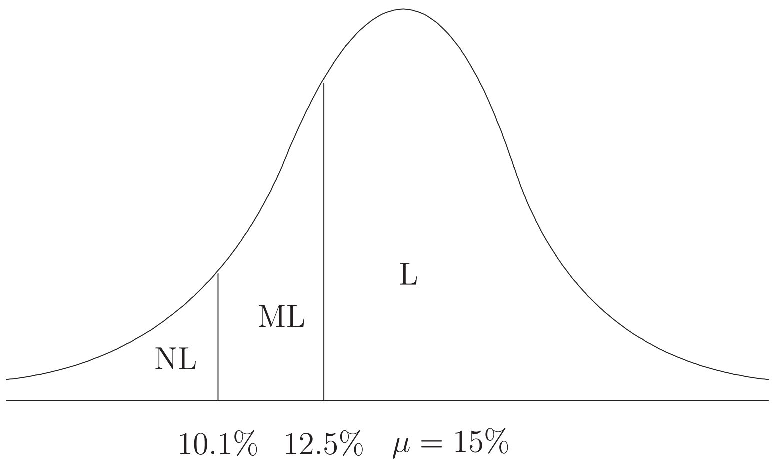


Figura 4.1:

–
$$\begin{aligned} P(L) =& P(X > 12.5)\\ =& 1 - P(X \leq 12.5)\\ =& 1 - P(Z \leq \frac{12.5-15}{2.5})\\ =& 1 - \Phi(-1)\\ =& 0.8413 \end{aligned}$$

–
$$\begin{aligned} P(NL) =& P(X < 10.1)\\ =& 1 - P(Z \leq \frac{10.1-15}{2.5})\\ =& \Phi(-1.96)\\ =& 0.025 \end{aligned}$$

–
$$\begin{aligned} P(ML) =& P(10.1 < X < 12.5)\\ =& \Phi(-1) - \Phi(-1.96)\\ =& 0.1337 \end{aligned}$$

Luego la cantidad de personas que caerán en cada categoría es

N_1 : número de personas en al categoría L=5400 $\times$ 0.8413=4543
N_2 : número de personas en al categoría NL=5400 $\times$0.025=135
N_3 : número de personas en al categoría ML=5400$\times$0.1337=722

b) $n = 5$; Y : número de personas que logran el objetivo en su totalidad,
$Y \sim Bin(n = 5, p = 0.8413)$

$\therefore P(Y = 3) = \binom{5}{3}0.8413^3 0.1587^2 = 10 \times 0.5955 \times 0.0252 = 0.15$

PROBLEMA 2[2]

El tiempo que un alumno de la carrera de Química y Farmacia, dedica diariamente a estudiar el curso de Métodos Estadísticos, es una variable aleatoria normal con media de 5 horas y desviación estándar de 2 horas.

a) Si el 25% de los alumnos estudian más de x horas diarias. Encuentre el valor de x.

Supongamos que cada padre le paga a un hijo por estudiar estas horas diarias y además definamos este pago mediante la función $P(X) = 1000X + 500$.

b) Verificar que $P(X)$ es una distribución normal, y encuentre μ_Y y σ_Y^2.

c) ¿Calcular la probabilidad de que un estudiante gane entre \$3500 y \$6200 ?

SOLUCIÓN

a) X : "tiempo que un alumno dedica en estudiar", donde $X \sim N(5, 4)$

$$\begin{aligned} P(X \geq x) &= 0.25 \\ 1 - P(X \leq x) &= 0.25 \\ P(Z \leq \tfrac{x-5}{2}) &= 0.75 \end{aligned}$$

luego $x = 6.3489796$

b) **Primera forma**

Como $X \sim N(5, 4)$, entonces por teorema visto en clase $P(X) = 1000X + 500 \sim N(1000 \times 5 + 500 = 5500, 1000^2 \times 4 = 4000000)$, aquí sus parámetros son $\mu_P = 5500$, $\sigma_P^2 = 4000000$

Segunda forma

$P = 1000X + 500$ lo que es equivalente $\frac{P-500}{1000} = x$

[2] I2 segundo semestre 2003

$$
\begin{aligned}
f_P(p) &= f_X\left(\tfrac{p-500}{1000}\right)\tfrac{1}{1000} \\
&= \frac{1}{\sqrt{(2\times\pi\times 4\times 1000^2)}}e^{-\frac{1}{2}\frac{(p-5500)^2}{4\times 1000^2}}
\end{aligned}
$$

Luego la función de densidad se distribuye $X \sim N(\mu_P = 5500, \sigma^2 = 4000000)$.

c)

$$
\begin{aligned}
P(3500 \leq p \leq 6200) &= P(-\tfrac{3500-5500}{2000} \leq z \leq \tfrac{6200-5500}{2000}) \\
&= P(-1 \leq z \leq 0.35) \\
&= \Phi(0.35) - \Phi(-1) \\
&= 0.6368 - 0.1587 \\
&= 0.4781
\end{aligned}
$$

La probabilidad de que un estudiante gane entre \$3500 y \$6200 es de 0.4781.

PROBLEMA 3[3]

El peso de un moderno zapato deportivo para correr tiene una distribución normal con media 12 onzas y desviación estándar de 0.5 onzas.

a) ¿Cuál es la probabilidad de que el zapato pese más de 13 onzas?

b) ¿Cuál debe ser la desviación estándar del peso para que la compañía que los produce pueda garantizar que el 99,9% de los zapatos tiene un peso menor que 13 onzas?

SOLUCIÓN

Sea X :"peso del zapato", $X \sim N(12, (0.5)^2)$

a) $P(X > 13) = P(Z > \frac{13-12}{0.5}) = P(Z > 2) = P(Z < -2) = 0.0228$

b) $P(X < 13) = 0.999 \Rightarrow P(Z < \frac{13-12}{\sigma}) = 0.999$

Luego $\frac{1}{\sigma} = 3.1$, entonces $\sigma = 0.322$

PROBLEMA 4[4]

Se tiene la siguiente función

$$
f(x) = \begin{cases} cx & 0 \leq x \leq 1 \\ c(2-x)/2 & 1 < x \leq 2 \\ 0 & e.o.c. \end{cases}
$$

[3] Examen segundo semestre de 2001
[4] I2 segundo semestre de 2003

Encuentre el valor de c de modo que la función f sea de densidad, obtenga la función de distribución acumulada y calcular $P(X \leq 1.5)$, $P(X = 1)$ y $P(0.5 \leq X \leq 1.5)$

<u>SOLUCIÓN</u>

Se debe cumplir que $\int f(x)dx = 1$ y $f(x) \geq 0$

$$\begin{array}{rcl} \int_0^1 cxdx + \int_1^2 \frac{c(2-x)}{2}dx & = & 1 \\ c(\frac{1}{2} + \frac{1}{4}) & = & 1 \\ c & = & \frac{4}{3} \end{array}$$

Luego

$$f(x) = \begin{cases} 4x/3 & 0 \leq x \leq 1 \\ 2(2-x)/2 & 1 < x \leq 2 \\ 0 & e.o.c. \end{cases}$$

Obtengamos la función de distribución acumulada

- Si $0 \leq t$, entonces $F_X(t) = 0$
- Si $0 \leq t < 1$, entonces

$$F_X(t) = \int_0^t \frac{4}{3}xdx = \frac{2}{3}t^2$$

- Si $1 \leq t < 2$, entonces

$$\begin{array}{rcl} F_X(t) & = & \int_0^1 \frac{4}{3}xdx + \int_1^t \frac{2}{3}(2-x)dx \\ & = & 1 - \frac{(2-t)^2}{3} \end{array}$$

- Si $t \geq 2$, entonces

$$F_X(t) = \int_0^1 \frac{4}{3}xdx + \int_1^2 \frac{2}{3}(2-x)dx + \int_2^t 0dx = 1$$

Luego se tiene la función de distribución acumulada

$$F_X(t) = \begin{cases} 0 & t \leq 0 \\ 2t^2/3 & 0 \leq t < 1 \\ 1 - (2-t)^2/3 & 1 \leq t < 2 \\ 1 & t \geq 2 \end{cases}$$

De aquí podemos obtener las probabilidades que se piden

- $P(X \leq 1.5) = F_X(1.5) = \frac{11}{12} = 0.9166667$

- $P(X = 1) = 0$, ya que F es continua en $X = 1$

- $$\begin{aligned} P(0.5 \leq X \leq 1.5) &= P(0.5 \leq X \leq 1.5) \\ &= P(X \leq 1.5) - P(X \leq 0.5) \\ &= F_X(1.5) - F_X(0.5) \\ &= 0.75 \end{aligned}$$

PROBLEMA 5[5]

Una amplia experiencia en ventiladores de cierto tipo, empleados en motores diesel, ha sugerido que la distribución exponencial es una buen modelo para el tiempo hasta que se presente una falla. Suponga que el tiempo medio hasta una falla es de 25 000 horas.

a) ¿Cuál es la probabilidad de que un ventilador seleccionado al azar dure por lo menos 20 000 horas? ¿a lo sumo 30 000 horas? y ¿entre 20 000 y 30 000 horas?

b) ¿Cuál es la probabilidad de que la duración de un ventilador exceda el valor medio en más de 2 desviaciones estándar? y ¿en más de 3 desviaciones estándar?

SOLUCIÓN

T :"tiempo hasta que se presenta una falla" $\quad T \sim Exp(\lambda)$

$\mu = E[T] = \frac{1}{\lambda} = 25000 \Rightarrow \lambda = \frac{1}{25000}$

$F_T(x) = 1 - e^{-\lambda x}$

$$\begin{aligned} V[T] &= \frac{1}{\lambda^2} \\ &= (25000)^2 \\ \Rightarrow \qquad \sigma &= 25000 \end{aligned}$$

a) –

$$\begin{aligned} P(T > 20000) &= 1 - P(T \leq 20000) \\ &= 1 - F_T(20000) \\ &= 1 - \left[1 - e^{-\frac{20000}{25000}}\right] \\ &= e^{-\frac{20}{25}} \\ &= 0.449 \end{aligned}$$

[5] TAV 2004

–

$$\begin{aligned}P(T \leq 30000) =&F_T(30000)\\ =&1 - e^{\frac{-30000}{25000}}\\ =&0.699\end{aligned}$$

Luego

$$\begin{aligned}P(20000 < T < 30000) =&P(T \leq 30000) - P(T \leq 20000)\\ =&0.699 - [1 - 0.449]\\ =&0.148\end{aligned}$$

b) –

$$\begin{aligned}P(T > \mu + 2\sigma) =&P(T > 25000 + 2(25000))\\ =&P(T > 75000)\\ =&1 - P(T \leq 75000)\\ =&1 - F_T(75000)\\ =&0.05\end{aligned}$$

–

$$\begin{aligned}P(T > \mu + 3\sigma) =&P(T > 25000 + 3(25000))\\ =&P(T > 100000)\\ =&1 - P(T \leq 100000)\\ =&1 - F_T(100000)\\ =&0.018\end{aligned}$$

PROBLEMA 6[6]

En una industria química, la venta mensual de cierto producto, en miles de libras, está representado por una v.a. X con función de densidad

$$f(x) = \begin{cases} x/4, & 0 \leq x < 2\\ (4-x)/4, & 2 \leq x < 4\\ 0, & \text{e.o.c.}\end{cases}$$

a) Comprobar que la función es de densidad.

b) Determinar la función de distribución acumulada de X y calcular $P(X = 2)$, $P(1.5 \leq X \leq 3.5)$.

[6]examen segundo semestre de 2003

c) Si se sabe que la venta en un mes dado no alcanza a 3000 libras, ¿cuál es la probabilidad que se haya tenido una venta de a lo menos 1500 libras?

d) Sea $Y = 2X - 3$. Determine $P(Y > 2)$.

SOLUCIÓN

a) Se debe cumplir que $\int f(x)dx = 1$ y $f(x) \geq 0, \forall x$

$$\int_0^2 \frac{x}{4}dx + \int_2^4 \frac{(4-x)}{4}dx = \frac{x^2}{8}\Big|_0^2 + \left(x - \frac{x^2}{8}\right)\Big|_2^4$$

$$= \frac{1}{2} + \left(2 - \frac{3}{2}\right)$$

$$= 1$$

b) Obtengamos la función de distribución acumulada

- Si $0 \leq t$, entonces $F_X(t) = 0$
- Si $0 \leq t < 2$, entonces

$$F_X(t) = \int_0^t \frac{x}{4}dx$$

$$= \frac{x^2}{8}\Big|_0^t$$

$$= \frac{t^2}{8}$$

- Si $2 \leq t < 4$, entonces

$$F_X(t) = \int_0^2 \frac{x}{4}dx + \int_2^t \frac{4-x}{4}dx$$

$$= \frac{x^2}{8}\Big|_0^2 + \left(x - \frac{x^2}{8}\right)\Big|_2^t$$

$$= \frac{1}{2} + t - \frac{t^2}{8} - 2 + \frac{1}{2}$$

$$= t - \frac{t^2}{8} - 1$$

– Si $t \geq 4$, entonces

$$F_X(t) = \int_0^2 \frac{x}{4}dx + \int_2^4 \frac{4-x}{4}dx$$

$$=1$$

Luego se tiene la función de distribución acumulada

$$F_X(t) = \begin{cases} 0 & t \leq 0 \\ t^2/8 & 0 \leq t < 2 \\ t - t^2/8 - 1 & 2 \leq t < 4 \\ 1 & t \geq 4 \end{cases}$$

De aquí podemos obtener las probabilidades que se piden

–

$$P(1,5 \leq X \leq 3.5) = F_X(3.5) - F_X(1.5)$$

$$= 0.97 - 0.28$$

$$= 0.69$$

– $P(X = 2) = 0$, ya que F es continua en $X = 2$

c)

$$P(X > 1.5|X < 3) = \frac{P(1.5 < X < 3)}{P(X < 3)}$$

$$= \frac{F_X(3) - F_X(1.5)}{F_X(3)}$$

$$= \frac{0.875 - 0.28}{0.875}$$

$$= 0.68$$

La probabilidad que haya tenido una venta de a lo menos 1500 libras, dado que la venta en un mes no alcanza a 3000 libras es de 0.68

d)

$$P(Y > 2) = P(2X - 3 > 2)$$

$$= P(X > 5/2)$$

$$= 1 - F_X(5/2)$$

$$= 0.28125$$

PROBLEMA 7[7]

Una envasadora de harina de trigo, es ajustada de tal manera que la probabilidad de que llene una bolsa con menos de 6.0204 kg. sea 0.102 y que la llene con más de 6.1984 kg. sea 0.008. Un comerciante que desea comprar este producto elige 10 bolsas de la producción de un día y decide comprar una gran partida, si a lo más una de estas bolsas pesan menos de 6.08 kg.

a) ¿Cuál es la probabilidad de que el comerciante compre la partida si es llenado de la envasadora se distribuye aproximadamente normal?

b) Si elige bolsas uno a uno de la producción diaria ¿cuál es la probabilidad de que encuentre la primera bolsa que pese menos de 6 kg. en la segunda elección?

SOLUCIÓN

[7] TAV 2004

X :Contenido de harina de trigo en cada bolas (kg); $X \sim N(\mu, \sigma^2)$

De los datos podemos obtener

$$\begin{array}{rcl} P(X \leq 6.0204) & = & 0.102 \\ P(X \leq 6.1984) & = & 0.992 \end{array}$$

$$\Rightarrow \quad \begin{array}{rcl} \frac{6.0204-\mu}{\sigma} & = & -1.27 \\ \frac{6.1984-\mu}{\sigma} & = & 2.41 \end{array}$$

Luego $\mu = 6.08$ (kg), $\sigma = 0.048$ (kg); $P(X < 6.08) = 0.5$

a) Y : "n° de bolsas que pesan a lo más 6.08 kg" $Y \sim Bin(10, 0.5)$

$$\begin{aligned} P(Y \leq 1) =& \binom{10}{0} 0.5^10 + \binom{10}{1} 0.5^10 \\ =& 0.0107 \end{aligned}$$

b) W :"n° de bolsas revisadas hasta encontrar la primera que pese menos de 6(kg)".

$W \sim Geom(p)$; $p = P(X < 6) = 0.0475$

$$\begin{aligned} P(W = 2) =& 0.0475 \times (1 - 0.0475) \\ =& 0.04524 \end{aligned}$$

PROBLEMA 8

El tiempo de reparación en (en horas) T de un artículo sigue aproximadamente una distribución de probabilidad con densidad

$$f(t) = te^{-t}, t > 0$$

a) El factor de depreciación Z se define por $Z = e^{-\alpha T}$. Calcular su media y varianza.

b) El costo de reparación de un artículo es $S + cT$, donde la constante c es un costo por unidad de tiempo y la variable aleatoria S toma los valores s_1 y s_2 con probabilidad p y $1 - p$ respectivamente. Calcule el costo esperado.

SOLUCIÓN

Sea T: tiempo de reparación de un artículo (en horas).

a) Sea $Z = e^{-\alpha T}$ "factor de depreciación"

$$E[Z] = \int_0^{\infty} e^{-\alpha t} t e^{-t} dt$$

$$= \int_0^{\infty} t e^{-(\alpha+1)t} dt$$

sea $s = (\alpha + 1)t$

$$E[Z] = \int_0^{\infty} \frac{s}{(\alpha+1)^2} e^{-s} ds$$

$$= \frac{1}{(\alpha+1)^2} \int_0^{\infty} s e^{-s} ds$$

$$= \frac{1}{(\alpha+1)^2}$$

$\therefore E[Z] = \frac{1}{(\alpha+1)^2}$

$V[Z] = E[Z^2] - (E[Z])^2$

$$E[Z^2] = \int_0^{\infty} e^{-2\alpha t} t e^{-t} dt$$

$$= \int_0^{\infty} t e^{-(2\alpha+1)t} dt$$

$$= \int_0^{\infty} \frac{s}{(2\alpha+1)^2} e^{-s} ds$$

$$= \frac{1}{(2\alpha+1)^2}$$

Luego $V[Z] = \frac{1}{(2\alpha+1)^2} - \frac{1}{(\alpha+1)^4}$

b) Costo de reparación de un artículo$=S + cT$

$$S = \begin{cases} s_1, & P(S = s_1) = p \\ s_2, & P(S = s_2) = 1 - p. \end{cases}$$

Costo esperado$=E(S+cT)=E(S)+cE[T]$

$$E[s] = s_1P(S=s_1)+s_2P(S=s_2)$$

$$= s_1p+s_2(1-p)$$

$$E[T] = \int_0^\infty t^2e^{-t}dt$$

$$= -t^2e^{-t}\Big|_0^\infty + 2\int_0^\infty te^{-t}dt$$

$$= 2\left(-te^{-t}\Big|_0^\infty + \int_0^\infty e^{-t}dt\right)$$

$$= -2e^{-t}\Big|_0^\infty$$

$$= 2$$

Luego $E[S+cT]=s_1p+s_2(1-p)+2c$.

PROBLEMA 9

Considere los tiempos de reparación $T_1,\ldots,T_n$ de n artículos como los descritos en el Problema 1. Estos tiempos se suponen independientes entre sí.

a) Verifique que el valor esperado del tiempo promedio de reparación de 32 items es 2 y que su desviación estándar es 0.25.

Aproxime la distribución del tiempo promedio de reparación por una distribución normal de media 2 y desviación estándar 0.25 y en base a esto:

b) Calcule $P(U_{32}<1.8)$.

c) Encuentre una cota superior c tal que $P(U_{32}>c)=0.90$.

d) Encuentre la probabilidad de que dos o más de estos 32 artículos tengan un tiempo de reparación inferior a una hora.

e) Sea Φ la función de distribución acumulada de la distribución $N(0.1)$. En términos de Φ encuentre una expresión aproximada para la probabilidad de que la mitad o menos de los 32 artículos tengan un tiempo de reparación inferior a 1.8 horas. (no se requiere valuar numéricamente esta expresión)

SOLUCIÓN

Sea $T_1, \ldots, T_n$ tiempos de recuperación de n artículos; los T_i son independientes entre sí; $i = 1, \ldots, n$

a)

$$E[\overline{T}] = E\left[\sum_{i=1}^{n} T_i/n\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n} E[T_i]$$

$$= E[T]$$

$$= E[T_j], \forall j$$

$$\begin{aligned} E[T_1] &= \int_0^\infty t^2 e^{-t} dt &= 2! &= 2 \\ E[T_2] &= \int_0^\infty t^2 t e^{-t} dt &= 3! &= 6 \end{aligned}$$

Así $V[T_1] = 6 - 4 = 2$

$$V[\overline{T}] = V\left[\sum_{i=1}^{n} T_i/n\right]$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} V[T_i], \text{ por independencia de los} T_i$$

$$= \frac{2}{n}$$

Para $n = 32$

$V[\overline{T}] = \frac{2}{32} = \frac{1}{16} \Rightarrow \sigma_{\overline{T}} = \frac{1}{\sqrt{16}} = 0.25$

b) Sea U_{32} : "tiempo promedio de reparación de 32 items"$U_{32} \sim N(2; (0.25)^2)$

$$P(U_{32} < 1.8) = P\left(\frac{U_{32} - \mu}{\sigma} < \frac{1.8 - 2}{0.25}\right)$$

$$= \Phi(-0.8)$$

$$= 0.2119$$

c)

$$P(U_{32} > c) = 0.90$$

$$P\left(\frac{U_{32} - \mu}{\sigma} > \frac{c - 2}{0.25}\right) = 0.90$$

$$\Phi\left(\frac{c - 2}{0.25}\right) = 0.1$$

$$c = 2 - 0.25 \times 1.28$$

d) Sea Y :"número de artículos con tiempo de reparación inferior a una hora $Y \sim Bin(32, p)$

$$p = P(T_1 \leq 1)$$

$$= \int_0 1te^{-t}dt$$

$$= -te^{-t}\Big|_0^1 - e^{-t}\Big|_0^1$$

$$= -e^{-1} - e^{-1} + 1$$

$$= 1 - 2e^{-1}$$

Entonces

$$P(Y \geq 2) = 1 - P(Y = 0) - P(y = 1)$$

$$= 1 - \binom{32}{0}\left(1 - \frac{2}{e}\right)^0 \left(\frac{2}{e}\right)^{32} - \binom{32}{1}\left(1 - \frac{2}{e}\right)\left(\frac{2}{e}\right)^{31}$$

$$= 1 - \left(\frac{2}{e}\right)^{32} - 32\left(1 - \frac{2}{e}\right)\left(\frac{2}{e}\right)^{31}$$

e) Sea Y :"número de artículos con tiempo de reparación inferior a 1.8 hrs"$Y \sim Bin(32, p)$, donde

$$p = P(T_1 < 1.8)$$

$$= 1 - e^{-1.8} - 1.8e^{-1.8}$$

$$= 1 - 2.8e^{-1.8}$$

La distribución aproximada de Y es:

$$Y \sim N(32p, 32p(1-p))$$

$$P(Y \leq 16) = P\left(Z \leq \frac{16 - 32p}{\sqrt{32p(1-p)}}\right)$$

$$= \Phi\left(\frac{16 - 32p}{\sqrt{32p(1-p)}}\right)$$

PROBLEMA 10

Suponga que el número de horas X que funcionará una máquina antes de fallar es una variable aleatoria con distribución Normal de parámetros $\mu = 720$ y $\sigma^2 = 48^2$.
Suponga que en el momento en que la máquina comienza a funcionar Ud. debe decidir cuando el inspector regresará a revisarla. Si el vuelve antes de que la máquina falle, se ocasiona un costo de a dólares por haber desperdiciado una inspección. Si vuelve después de que la máquina haya fallado, se ocasiona un costo de b dólares por el no funcionamiento de la máquina.

a) Determine una expresión para el costo esperado, considerando que el tiempo hasta que el inspector vuelve a inspeccionar la máquina es t horas.

b) Suponga que el inspector decide volver en un tiempo de $t = 816$ hrs. Calcule la probabilidad de que el inspector llegue tarde a la inspección, es decir, la máquina ya ha dejado de funcionar.

c) Se observa este proceso durante 15 períodos. Determine la probabilidad de que el inspector llegue tarde más de 12 veces.

SOLUCIÓN

X :"tiempo de funcionamiento de una máquina hasta que falla", $X \sim N(720, 48^2)$

a)

$$Costo = \begin{cases} a & x > t \\ b & x < t \end{cases}$$

$$\begin{aligned} E(Costo) &= aP(X > t) + bP(X < t) \\ &= a - aP(X < t) + bP(X < t) \\ &= a + (b - a)F_Z\left(\frac{t - 720}{48}\right) \end{aligned}$$

b)

$$\begin{aligned} P(X < 816) &= P\left(\frac{X - 720}{48} < \frac{816 - 720}{48}\right) \\ &= P\left(Z < 2\right) \\ &= 0.9772 \end{aligned}$$

c) X :número de veces que el inspector llega tarde, $X \sim Bin(15, 0.9772)$

$$P(X > 12) = \sum_{13}^{15} \binom{15}{x} (0.9772)^x (0.0228)^{15-x}$$

PROBLEMA 11[8]

Suponga que $Y_1, \ldots, Y_{40}$ es una muestra aleatoria de mediciones con respecto a las proporciones de impurezas en unas muestras de minerales de hierro. Cada Y_i Tiene una función de densidad de probabilidad dada por

$$f_Y(y) = \begin{cases} 3y^2 & 0 \leq y \leq 1 \\ 0 & e.o.c. \end{cases}$$

Un comprador potencial rechazará el mineral si $\overline{Y}$ es mayor que 0.7. Calcular $P(\overline{Y} > 0.7)$ para un tamaño de muestra igual a 40.

SOLUCIÓN

$$\begin{aligned} P(\overline{Y} > 0.7) &= 1 - P(\overline{Y} \leq 0.7) \\ &= 1 - P\left(\tfrac{\overline{Y}-\mu}{\sigma/n} \leq \tfrac{0.7-\mu}{\sigma/n}\right) \\ &= 1 - P\left(Z \leq \tfrac{0.7-\mu}{\sigma/n}\right) \\ &= 1 - \Phi\left(\tfrac{0.7-\mu}{\sigma/n}\right) \qquad (*) \end{aligned}$$

Nos falta calcular μ y σ^2

- $$\begin{aligned} \mu &= E[Y] \\ &= \int_0^1 y3y^2 dy \\ &= \left.\tfrac{3y^4}{4}\right|_0^1 \\ &= \tfrac{3}{4} \end{aligned}$$

- $$\begin{aligned} \sigma^2 &= E[Y^2] - (E[Y])^2 \\ &= \int_0^1 y^2 3y^2 dy - \left(\tfrac{3}{4}\right)^2 \\ &= \left.\tfrac{3y^5}{5}\right|_0^1 - \tfrac{9}{16} \\ &= \tfrac{3}{5} - \tfrac{9}{16} \\ &= \tfrac{3}{80} \end{aligned}$$

[8] TAV 2004

Volviendo a $(*)$ tenemos

$$\begin{aligned} P(\overline{Y} > 0.7) &= 1 - \Phi\left(\frac{0.7-3/4}{\sqrt{3/80}/\sqrt{40}}\right) \\ &= 1 - \Phi(-1.63) \\ &= 0.9484 \end{aligned}$$

PROBLEMA 12[9]

Supongamos que extraemos una muestra simple(aleatoria) de una población que se distribuye según una uniforme $U(0,2)$. Encontrar aproximadamente la probabilidad de que la media muestral se encuentre entre 0.8 y 1.1, si el tamaño de muestra es 48.

SOLUCIÓN

Se tiene que $X_i \sim U(0,2)$, $i = 1, \ldots, n$, de aquí obtenemos

$E[X_i] = \frac{2}{2} = 1 \qquad V[X_i] = \frac{1}{3} \qquad , i = 1, \ldots, n$

Luego

- $$\begin{aligned} E[\overline{X}_{48}] &= \frac{1}{48} E\left[\sum_{i=1}^{48} X_i\right] \\ &= 1 \end{aligned}$$

- $$\begin{aligned} V[\overline{X}_{48}] &= \left(\frac{1}{48}\right)^2 V\left[\sum_{i=1}^{48} X_i\right] \\ &= \frac{48}{48^2} V[X_i] \\ &= \frac{1}{48} \times \frac{1}{3} \\ &= \frac{1}{144} \end{aligned}$$

Se nos pide $P(0.8 \leq \overline{X}_{48} \leq 1.1)$, y utilizando el $T.C.L.$ tenemos

$$\begin{aligned} P(0.8 \leq \overline{X}_{48} \leq 1.1) &= \Phi\left(\frac{1.1-1}{1/12}\right) - \Phi\left(\frac{0.8-1}{1/12}\right) \\ &= \Phi(1.2) - \Phi(-2.4) \\ &= 0.8767 \end{aligned}$$

[9]Examen segundo semestre de 2003

4.2 Ejercicios Propuestos

Distribuciones Continuas

1. [10] Considere que X es la temperatura a la que tiene lugar cierta reacción química y suponga que X tiene densidad

$$f(x) = \begin{cases} \frac{1}{9}(4 - x^2), & -1 \leq x \leq 1 \\ 0, & \text{en otro caso.} \end{cases}$$

   a) Construya la gráfica de $f(x)$.
   b) Determine la función de distribución acumulada y trácela.
   c) ¿Es 0 la temperatura mediana en la cual tiene lugar la reacción? Si no es así, ¿es la temperatura mediana menor o mayor que 0?
   d) Suponga que está reacción se realiza independientemente una vez en cada uno de los diferentes laboratorios y que la pdf del tiempo de reacción en cada laboratorio es la de arriba. Sea Y=número entre los diez laboratorios y que la pdf del tiempo de reacción en cada laboratorio es la de arriba. Sea Y=número entre los diez laboratorios en los que la temperatura rebasa 1. ¿Qué clase de distribución tiene Y? (Dé el nombre y valores de cualesquiera parámetros).

2. [11] El artículo " Computer Assisted Net Weight Control"(*Quality Progress*, 1983, pp. 22-25) sugiere una distribución normal, con media de 137.2 onza y desviación estándard de 1.6 onzas, para el contenido real de frascos de cierto tipo. El contenido establecido era 135 onzas.

   a) ¿Cuál es la probabilidad de que un solo frasco contenga más que el contenido establecido?
   b) Entre diez frascos seleccionados al azar, ¿cuál es la probabilidad de que por lo menos ocho contengan más del contenido establecido?
   c) Si se supone que la media permanece en 137.2, ¿a qué valor tendría que haberse cambiado la desviación estándard para que 95% de todos los frascos contengan más de lo establecido?

3. [12] La demanda semanal de gas propano (en miles de galones) de una distribución en particular es una v.a. X con densidad

$$f(x) = \begin{cases} 2\left(1 - \frac{1}{x^2}\right), & 1 \leq x \leq 2 \\ 0, & \text{en otro caso.} \end{cases}$$

   a) Calcule la función de distribución acumulada de X.
   b) Obtenga una expresión para el (100p)mo percentil.¿Cuál es el valor de $\tilde{\mu}$?

[10] I2 segundo semestre de 2002
[11] I2 segundo semestre de 2002
[12] I2 segundo semestre de 2002

c) Calcule $E(X)$ y $V(X)$.

d) Si 1500 galones están en existencia al principio de semana y no se recibe nuevo suministro durante la semana, ¿cuánto de los 1500 galones se espera que queden al fin de semana? [Sugerencia: sea $h(x)$=cantidad que queda cuando la demanda es $= x$.]

4. El espesor de la capa de sustancia fotoprotectora que se aplica a las obleas en el proceso de fabricación de semiconductores en cierta área de la oblea, tiene una distribución uniforme entre 0.2050 y 0.2150 micrómetros. Obtenga la función de distribución acumulada del espesor de la sustancia fotoprotectora.

   a) Calcule la proporción de obleas en las que el espesor de la sustancia es mayor que 0.2125 micrómetros.

   b) ¿Qué espesor exceden el 10% de las obleas?

   c) Calcule la media y la varianza del espesor de la sustancia fotoprotectora.
   Resp:a)0.25 b)0.214 c)0.21, 8.33×10^{-6}.

5. La destiladora Los Vascos produce entre 200 y 300 galones de vino diarios. La distribución uniforme es la que mejor describe este proceso.

   a) ¿Cuánto vino se produce al día en promedio?

   b) ¿Cuál es la cantidad de variabilidad en el número de galones de vino producidos de un día a otro?

   c) ¿En qué porcentaje de los días puede esperarse que la producción caiga entre 220 y 270 galones?

   d) ¿Cuál es la probabilidad de que la producción de mañana sea mayor que 280 galones?
   Resp: a)250 b)28.9 c)0.5 d)0.2.

6. El tiempo que transcurre entre llamadas a una empresa de artículos computacionales tiene una distribución exponencial con un tiempo promedio entre llamadas en un lapso de 15 minutos.

   a) ¿Cuál es la probabilidad de que no haya llamadas en un lapso de 30 minutos?

   b) ¿Cuál es la probabilidad de recibir la primera llamada entre 5 y 10 minutos después de haber abierto la empresa?

   c) Calcule la dimensión de un intervalo de tiempo, de modo tal que la probabilidad de recibir al menos una llamada en ese período sea 0.9.
   Resp: a)0.1353 b)0.2031 c)34.54 min.

7. El tiempo (en horas) requeridas para reparar una máquina es una distribución exponencial $\lambda = 1/2$:

   a) ¿Cuál es la probabilidad que el tiempo de reparación exceda 2 horas?

b) ¿Cuál es la probabilidad que una reparación tome al menos 10 horas dado que su duración exceda en 9 horas?
Resp: a)0.5128 b)0.6065

8. El tiempo de vida de los reguladores de voltaje de los automóviles tiene una distribución exponencial con un tiempo de vida medio de seis años. Una persona compra un automóvil que tiene una antigüedad de seis años, con un regulador en funcionamiento, y planea tenerlo por espacio de seis años.

   a) ¿Cuál es la probabilidad de que el regulador de voltaje falle en el lapso de seis años?

   b) Si el regulador falla después de tres años de haber efectuado la compra del automóvil y se reemplaza,¿cuál es el tiempo promedio que transcurrirá hasta que el regulador vuelva a fallar?

9. Los accidentes de automóviles ocurren en Santiago, durante un fin de semana largo(72 horas), según un proceso de Poisson a una tasa de 10 por hora. Determinar la probabilidad que el segundo accidente ocurra después de una hora.
Resp: $11\exp(-10)$.

10. En el Restaurante Los Buenos Muchachos, se ha instalado una máquina para la venta de cervezas. La máquina puede regularse de modo que la cantidad media de cerveza por vaso sea la que se desee; sin embargo, en cualquier caso esta cantidad es una variable aleatoria con una distribución normal con una desviación estándar de 5.9cc.

    a) Si el nivel se ajusta a 501 cc,¿que porcentaje de los vasos contendrá menos de 487 cc?

    b) ¿A que nivel medio debe ajustarse la máquina para el 83.15% de los vasos contenga más de 490 cc?
    Resp:a)0.89% b)496

11. Suponga que los conteos registrados por un contador Gieger siguen un proceso Poisson con un promedio de dos conteos por minuto.

    a) ¿Cuál es el tiempo promedio entre conteos?

    b) ¿Cuál es la desviación estándar del tiempo entre conteos?

    c) Calcule x, de modo tal que la probabilidad de que ocurra por lo menos un conteo antes de x minutos, sea 0.95
    Resp:a)0.5 b)0.5 c)1.5

12. Si el volumen de una máquina automática en latas de una bebida gaseosa tiene una distribución normal con media de 12.4 onzas de líquido y desviación estándar de 0.1 onzas de líquido. Se pide:

    a) Defina la variable en estudio, y escriba su f.d.p.

    b) Interprete la $E[X]$.

c) Si se desechan todas las latas que tienen menos de 12.2 o más de 12.8 onzas de líquido. ¿Cuál es el porcentaje de latas desechadas?

d) Si la media de la operación de llenado puede ajustarse con facilidad, pero la desviación estándar sigue teniendo el mismo valor, 0.1 onzas de líquido,¿Qué valor debe darse a la media para que el 95.85% de todas las latas contengan más de 12.2 onzas de líquido?

13. La resistencia a la comprensión de una serie de muestras de cemento puede modelarse con una distribución normal con media 6000 kilogramos por centímetro cuadrado, y una desviación estándar de 100 kilogramo por centímetro cuadrado.

    a) ¿Cuál es la probabilidad de que la resistencia de una muestra sea menor que 6250 kg/cm^2

    b) ¿Cuál es la probabilidad de que la resistencia de una muestra se encuentre entre 5800 y 5900 kg/cm^2

    c) ¿Cuál es el valor de resistencia que excede el 95% de las muestras?
    Resp:a)0.99379 b)0.13595 c)5835.5 kg/cm^2

14. Suponga que el diámetro de un cable eléctrico está normalmente distribuido con un promedio de 0.8 pulgadas y una varianza de 0.004 pulgadas.

    a) Si se elige un cable al azar, ¿cuál es la probabilidad que su diámetro sea menor que 0.85 pulgadas?

    b) Qué diámetro deberían tener los cables de modo que el 87.7% de ellos no excedan de este valor?
    Resp: b)0.8232 pulgadas.

15. El tiempo X (minutos) para que un asistente de laboratorio prepare el equipo para un experimento tiene una distribución uniforme con A=25 y B=35.

    a) Escriba la fdp de X y trace su gráfica.

    b) ¿Cuál es la probabilidad de que el tiempo de preparación exceda de 33 min?

    c) ¿Cuál es la probabilidad de que el tiempo de preparación se encuentre a una distancia de 2 min del tiempo medio [*Sugerencia:* identifique μ de la gráfica de $f(x)$.]

    d) Para cualquier a tal que $25 < a < a + 2 < 35$, ¿cuál es la probabilidad de que el tiempo de preparación esté entre a y $a + 2$ min?

16. Sea X la distancia en metros que un animal se mueve desde su lugar de nacimiento hasta el primer territorio vacante que encuentra. Suponga que para las ratas canguro, X tiene una distribución exponencial con parámetro $\lambda = 0.01386$ (como lo sugiere el artículo "Competition and Dispersal from Multiple Nests", Ecology, 1997, pp.873-883)

    a) ¿Cuál es la probabilidad de que la distancia sea a lo sumo 100 m? ¿Cuando mucho 200 m? ¿Entre 100 y 200 m?

b) ¿Cuál es la probabilidad de que la distancia sea mayor que la distancia promedio en más de 2 desviaciones estándard?

c) ¿Cuál es el valor de la mediana de la distancia?

17. La dureza Rockwell de un metal se determina al golpear con un punto acerado (herramienta) la superficie del metal y después medir la profundidad de penetración del punto. Suponga que la dureza Rockwell de cierta aleación está normalmente distribuida con media de 70 y desviación estándard de 3 (la dureza Rockwell se mide en escala continua).

a) Si un espécimen es aceptable sólo si su dureza está entre 67 y 75, ¿cuál es la probabilidad de que un espécimen seleccionado al azar tenga una dureza aceptable ?

b) Si la escala aceptable de dureza es (70-c,70+c), ¿para qué valor de c tendría una dureza aceptable 95 % de todos los especímenes?

c) Si la escala aceptable es como el inciso a) y la dureza de cada diez especímenes es independiente, ¿cuál es el número esperado de especímenes aceptables entre los diez?

d) ¿Cuál es la probabilidad de que a lo sumo ocho de diez especímenes seleccionados independientemente tengan una dureza menor de 73.84? (*Sugerencia:* Y=número entre los diez especímenes con dureza menor de 73.84 es una variable binomial; ¿cuál es p?)

18. Para trasladarme al trabajo, primero debo abordar un autobús cerca de casa y después transbordar otro. Si el tiempo de espera (en minutos) en cada parada tiene una distribución uniforme con A=0 y B=5, entonces se puede demostrar que mi tiempo total de espera Y tiene la fdp

$$f(y) = \begin{cases} \frac{1}{25}y & 0 \leq y < 5 \\ \frac{2}{5} - \frac{1}{25}y & 5 \leq y \leq 10 \\ 0 & y < 0 \text{ o } y > 10. \end{cases}$$

a) Trace la gráfica de la pdf de Y.

b) Verifique que $\int_{-\infty}^{\infty} f(y)\partial y = 1$.

c) ¿Cuál es la probabilidad de que el tiempo total de espera sea a lo sumo de 3 min?

d) ¿Cuál es la probabilidad de que el tiempo total de espera sea a lo sumo de 8 min?

e) ¿Cuál es la probabilidad de que el tiempo total de espera esté entre 3 y 8 min?

f) ¿Cuál es la probabilidad de que el tiempo total de espera sea menos de 2 min o más de 6?

19. Sea X la cantidad de espacio ocupado por un artículo colocado en una caja de empaque de 1 pie^3. La fdp de X es

$$f(x) = \begin{cases} 90x^8(1-x) & 0 < x < 1 \\ 0 & \text{en otro caso.} \end{cases}$$

a) Grafique la fdp. A continuación determine la función de distribución acumulada de X y grafíquela.

b) ¿Cuánto es $P(X \leq 0.5)$, [es decir, $F(0.5)$]?

c) De acuerdo con el inciso a), ¿cuánto es $P(0.25 < X \leq 0.5)$?

d) ¿Cuál es el 75to percentil de la distribución?

e) Calcule $E(X)$ y σ_X.

f) ¿Cuál es la probabilidad de que X esté a menos de 1 desviación estándard de su valor medio?

Resp: a) $F(x) = \begin{cases} 0, & \text{para } x \leq 0; \\ 90[\frac{x^9}{9} - \frac{x^{10}}{10}], & \text{para } 0 < x < 1; \\ 1, & \text{para } x \geq 1. \end{cases}$
b) 0.0107 c) 0.0107,0.0107 d)0.9036 e)0.818, 0.111 f)0.839

20. Sea X la temperatura en ^{o}C en la cual tiene lugar cierta reacción química, y sea Y la temperatura en ^{o}F (esto es $Y = 1.8X + 32$).

    a) Si la mediana de la distribución de X es $\tilde{\mu}$, demuestre que $1.8\tilde{\mu}+32$ es la mediana de la distribución de Y.

    b) ¿Cómo está relacionado el 90mo percentil de la distribución de Y con el 90mo percentil de la distribución de X? Verifique su respuesta.

    c) Generalmente, si $Y = aX + b$, ¿cómo se relaciona cualquier percentil particular de la distribución de Y con el correspondiente percentil de la distribución de X?

    Resp: b) 1.8(90mo percentil para X)+32 c)$a(Xpercentil) + b$

21. Sea X=tiempo entre dos llegadas sucesivas en la ventanilla de atención de un banco local. Si X tiene una distribución exponencial con $\lambda = 1$ (que es idéntica a una distribución gamma estándard con $\alpha = 1$), calcule lo siguiente:

    a) El tiempo esperado entre dos llegadas sucesivas.

    b) La desviación estándar del tiempo entre llegadas sucesivas.

    c) $P(X \leq 4)$.

    d) $P(2 \leq X \leq 5)$.

22. En cada caso, determine el valor de la constante c que exprese correctamente el enunciado de la probabilidad.

    a) $\Phi(c) = 0.9838$

    b) $P(0 \leq Z \leq c) = 0.291$

    c) $P(Z \geq c) = 0.121$

d) $P(-c \leq Z \leq c) = 0.668$

e) $P(|Z| \geq c) = 0.016$

Resp: a) 2.14 b) 0.81 c) 1.17 d) 0.97 e) 2.41

23. Suponga que el diámetro de los árboles de determinado tipo, a la altura del pecho, se distribuye normalmente con media $\mu = 8.8$ y $\sigma = 2.8$, como se sugiere en el artículo "Simulating a Harvester-Forwarder Softwood Thinning"(*Forest Products J.*,mayo de 1997, pp. 36-41)

    a) ¿Cuál es la probabilidad de que el diámetro de un árbol, seleccionado al azar, sea a lo sumo 10 pulg? y ¿que sea mayor de 10 pulg?

    b) ¿Cuál es la probabilidad de que el diámetro de un árbol, seleccionado al azar, sea mayor de 20 pulg?

    c) ¿Cuál es la probabilidad de que el diámetro de un árbol seleccionado al azar este entre 5 y 10 pulg?

    d) ¿Que valor de c es tal que el intervalo (8.8-c, 8.8+c) incluya 98% de todos los valores de diámetro?

    Resp: a) 0.3336 b) Aproximadamente 0 c) 0.5795 d) 6.524

24. Suponga que sólo 40% de todos los automovilistas de cierto estado, usan con regularidad su cinturón de seguridad. Se selecciona al azar una muestra de 500 automovilistas. ¿Cuál es la probabilidad de que:

    a) entre 180 y 230 (inclusive) de los automovilistas de la muestra utilice su cinturón con regularidad?

    b) menos de 175 de los de la muestra utilicen su cinturón con regularidad? y ¿menos de 150?

    Resp: a) 0.9666 b) 0.0099, 0

25. Hay 40 estudiantes en un curso de estadística elemental. Con base en los años de experiencia, el instructor sabe que el tiempo necesario para calificar un primer examen seleccionado al azar, es una variable aleatoria con valor esperado de 6 min y desviación estándar de 6 min.

    a) Si los tiempos para calificar son independientes y el instructor comienza a calificar a las 6:50 p.m. y lo hace en forma continua, ¿cuál es la probabilidad (aproximada) de que termine de calificar antes del inicio de las noticias de las 11:00 p.m. por TV?

    b) ¿Si la sección deportiva empieza a las 11:10, ¿cuál es la probabilidad de que se pierda parte de esa sección si espera hasta terminar antes de encender el televisor?

Resp: a) 0.6026 b)0.2981

26. El tiempo utilizado por un solicitante seleccionado al azar, para llenar cierta forma de hipoteca, tiene una distribución normal con valor medio de 10 minutos y desviación estándar de 2 min. Si cinco individuos llenan una forma en un día y seis en otro, ¿cuál es la probabilidad de que la cantidad de tiempo promedio de la muestra diaria sea a lo sumo de 11 minutos?
Resp: 0.7720

Capítulo 5

Estimación

5.1 Ejercicios Resueltos

PROBLEMA 1[1]

Sean $X_1, X_2, \ldots, X_n$ una m.a. de tamaño n de una población con la siguiente función de densidad:

$$f(x_i) = \begin{cases} \frac{1}{x_i \sigma \sqrt{2\pi}} \exp\left\{-\frac{1}{2}\frac{(\ln(x_i)-\mu)^2}{\sigma^2}\right\}, & x_i > 0, i = 1, \ldots, n; \\ 0, & \text{e.o.c.} \end{cases}$$

a) Encontrar el EMV de μ , con σ^2 conocido.

b) Si $n = 3$ y $X_1 = e$, $X_2 = e^2$, $X_3 = e^3$. Evaluar μ_{EMV} encontrado en (a).

SOLUCIÓN

a) Sea

$$\begin{aligned} f(x) &= \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right\} \\ L &= \prod_{i=1}^{n} \frac{1}{x_i \sigma \sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\ln x_i - \mu}{\sigma}\right)^2\right\} \\ &= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \frac{1}{\prod_{i=1}^{n} x_i} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{\ln x_i - \mu}{\sigma}\right)^2\right\} \end{aligned}$$

Aplicando logaritmo a L se tiene la log-verosimilitud

[1] I3 segundo semestre de 2000

$$\begin{aligned}
\ln L &= -n\ln(\sigma\sqrt{2\pi}) - \ln\left(\prod_{i=1}^{n} x_i\right) - \tfrac{1}{2}\sum_{i=1}^{n}\left(\frac{\ln x_i - \mu}{\sigma}\right)^2 \\
\Rightarrow \quad \tfrac{\partial \ln L}{\partial \mu} &= \tfrac{1}{\sigma}\sum_{i=1}^{n}\left(\frac{\ln x_i - \mu}{\sigma}\right)
\end{aligned}$$

Luego

$$\begin{aligned}
\tfrac{\partial \ln L}{\partial \mu} &= 0 \\
\Rightarrow \quad \tfrac{1}{\sigma}\sum_{i=1}^{n}\left(\frac{\ln x_i - \mu}{\sigma}\right)^2 &= 0 \\
\Rightarrow \qquad\qquad \widehat{\mu} &= \frac{\sum\limits_{i=1}^{n} \ln x_i}{n}
\end{aligned}$$

b) Como $X_1 = e$, $X_2 = e^2$, $X_3 = e^3$, entonces $\ln X_1 = 1$, $\ln X_2 = 2$, $\ln X_3 = 3$

Luego

$$\begin{aligned}
\widehat{\mu} &= \frac{\sum\limits_{i=1}^{3} \ln x_i}{3} \\
&= \tfrac{1+2+3}{3} \\
&= 2
\end{aligned}$$

PROBLEMA 2[2]

Suponga que X sigue una distribución de Pareto, su función de densidad esta dada por:

$f(x|\alpha,\theta) = \theta\alpha^{\theta}x^{-\theta-1}, x \geq \alpha$ y $\theta \geq 1$. Asuma que $\alpha > 0$ es conocido y que $X_1, \ldots, X_n$ son v.a iid .

a) Encuentre un estimador de momentos para θ.

b) Determine el EMV de θ.

[2] I3 segundo semestre de 2003

<u>SOLUCIÓN</u>

a) Primero obtengamos el valor esperado de la distribución de Pareto

$$
\begin{aligned}
E[X] = & \int_0^\infty x\theta\alpha^\theta x^{-\theta-1}dx \\
= & \int_0^\infty \theta\alpha^\theta x^{-\theta}dx \\
= & \theta\alpha^\theta \int_0^\infty x^{-\theta}dx \\
= & \theta\alpha^\theta \frac{x^{-\theta+1}}{-\theta+1}\Big|_\alpha^\infty \\
= & \theta\alpha^\theta \frac{-\alpha^{-\theta+1}}{-\theta+1} \\
= & \frac{-\theta\alpha}{-\theta+1} \\
= & \frac{\theta\alpha}{\theta-1}
\end{aligned}
$$

luego el primer momento poblacional es $\mu_1 = E[X] = \frac{\theta\alpha}{\theta-1}$. Ahora despejando θ se tiene $\theta = h(\mu_1) = \frac{\mu_1}{\mu_1-\alpha}$ y además se tiene que el primer momento muestral es $M_1 = \frac{1}{n}\sum x_i = \overline{x}$, por lo tanto el estimador por momentos de θ es

$$
\begin{aligned}
\widehat{\theta} = & \frac{M_1}{M_1-\alpha} \\
= & \frac{\overline{x}}{\overline{x}-\alpha}
\end{aligned}
$$

b) Determine el EMV de θ

La función de verosimilitud L está dado por

$$\begin{aligned} L(\theta|x_1,\ldots,x_n) &= \prod_{i=1}^{n} f(x_i) \\ &= \theta^n \alpha^{n\theta} \left(\prod_{i=1}^{n} x_i \right)^{-\theta-1} \end{aligned}$$

Así la $\log -verosimilitud$

$$l = \ln L(\theta|x_1,\ldots,x_n) = n\ln\theta + n\theta\ln\alpha - (\theta+1)\ln\left(\prod_{i=1}^{n} x_i\right)$$

$$\begin{aligned} \frac{\partial l}{\partial \theta} &= \tfrac{n}{\theta} + n\ln\alpha - \ln\left(\prod_{i=1}^{n} x_i\right) &= 0 \\ \Rightarrow \tfrac{n}{\theta} &= \ln\alpha - \ln\left(\prod_{i=1}^{n} x_i\right) - n\ln\alpha \\ \Rightarrow \widehat{\theta} &= \frac{n}{\ln\left(\prod_{i=1}^{n} x_i\right) - n\ln\alpha} \end{aligned}$$

PROBLEMA 3[3]

Represente con X la proporción de tiempo asignado que un estudiante seleccionado al azar emplea trabajando en cierta prueba de aptitud, y suponga que la función de probabilidad de X es

$$f_X(x;\theta) = \begin{cases} (\theta+1)x^{\theta} & 0 \le x \le 1 \\ 0 & \text{e.o.c.} \end{cases}$$

donde $-1 < \theta$. Una muestra aleatoria de diez estudiante produce la siguiente información: $x_1 = 0.92$, $x_2 = 0.79$, $x_3 = 0.9$, $x_4 = 0.65$, $x_5 = 0.86$, $x_6 = 0.47$, $x_7 = 0.73$, $x_8 = 0.97$, $x_9 = 0.94$ y $x_{10} = 0.77$. Obtenga el estimador de máxima verosimilitud de θ y después calcule la estimación para los datos proporcionados.

[3] I2 TAV 2004

SOLUCIÓN

$$L(\theta|x_1,\ldots,x_n) = \prod_{i=1}^{n} f(x_i;\theta)$$

$$= \prod_{i=1}^{n}(\theta+1)x_i^{\theta}$$

$$= (\theta+1)^n\left(\prod_{i=1}^{n} x_i\right)^{\theta}$$

Apliquemos logaritmo a lo anterior

$$l(\theta) = \ln L(\theta|x_1,\ldots,x_n)$$

$$= n\ln(\theta+1) + \theta\sum_{i=1}^{n}\ln x_i$$

Ahora encontremos el máximo

$$\Rightarrow \quad \frac{\partial l}{\partial \theta} = \tfrac{n}{\theta+1} + \sum_{i=1}^{n}\ln x_i$$

$$\Rightarrow \quad \left.\frac{\partial l}{\partial \theta}\right|_{\theta=\widehat{\theta}} = 0$$

$$\Rightarrow \quad \tfrac{n}{\widehat{\theta}+1} + \sum_{i=1}^{n}\ln x_i = 0$$

$$\Rightarrow \quad \widehat{\theta} = -\frac{n}{\sum_{i=1}^{n}\ln x_i} - 1$$

La estimación por los datos proporcionados $\widehat{\theta} = -\frac{10}{-2.429503} - 1 = 3.116$

PROBLEMA 4[4]

Si X tiene distribución Gamma con parámetros $\alpha > 0$ y $\beta > 0$,($X \sim G(\alpha,\beta)$), entonces se función de densidad es de la forma

$$f(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)}, \qquad 0 < x < \infty$$

Considere α conocido, calcular el EMV de β.

SOLUCIÓN

$$\begin{aligned}
L(\beta|x_1,\ldots,x_n) &= \prod_{i=1}^{n} f(x_i;\beta) \\
&= \prod_{i=1}^{n} \frac{x_i^{\alpha-1}e^{-x_i/\beta}}{\beta^{\alpha}\Gamma(\alpha)} \\
&= \left(\prod_{i=1}^{n} x_i^{\alpha-1}\right) e^{-\sum_{i=1}^{n} x_i/\beta} \frac{1}{\prod_{i=1}^{n} \beta^{\alpha}\Gamma(\alpha)} \\
&= \left(\prod_{i=1}^{n} x_i^{\alpha-1}\right) \frac{e^{-\sum_{i=1}^{n} x_i/\beta}}{\beta^{n\alpha}(\Gamma(\alpha))^n}
\end{aligned}$$

Apliquemos logaritmo a lo anterior

$$\begin{aligned}
l(\beta) &= \ln L(\beta|x_1,\ldots,x_n) \\
&= \ln\left(\prod_{i=1}^{n} x_i^{\alpha-1}\right) - \frac{\sum_{i=1}^{n} x_i}{\beta} - n\beta\ln\beta - n\ln\Gamma(\alpha)
\end{aligned}$$

Ahora encontremos el máximo

[4] Examen segundo semestre de 2003

$$\begin{aligned}
\Rightarrow \qquad \frac{\partial l}{\partial \beta} &= \frac{\sum_{i=1}^{n} x_i}{\beta^2} - \frac{n\alpha}{\beta} \\
\Rightarrow \qquad \left.\frac{\partial l}{\partial \beta}\right|_{\beta=\widehat{\beta}} &= 0 \\
\Rightarrow \qquad \frac{\sum_{i=1}^{n} x_i}{\beta^2} - \frac{n\alpha}{\beta} &= 0 \\
\Rightarrow \qquad \widehat{\beta} &= \frac{\sum_{i=1}^{n} x_i}{n\alpha}
\end{aligned}$$

PROBLEMA 5

Sea X_1, X_2 una muestra aleatoria de tamaño 2 de X con distribución exponencial de parámetro λ desconocido. Consideremos a $\widehat{\theta_1} = \overline{X}$ y a $\widehat{\theta_2} = \sqrt{X_1 X_2}$ estimadores de $\mu = \frac{1}{\lambda}$. En términos del error cuadrático medio, ¿cuál de los dos es mejor?

Para este problema será util considerar $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$; $\quad \Gamma(1/2) = \sqrt{\pi}$
y $E(\sqrt{X_1 X_2}) = E(\sqrt{X_1})E(\sqrt{X_2})$

SOLUCIÓN

El $ECM[\widehat{\theta_1}] = V[\widehat{\theta_1}] + B^2$

Veamos si $\widehat{\theta_1} = \overline{X}$ es insesgado respecto μ

$$\begin{aligned}
E[\overline{X}] &= E\left[\frac{X_1 + X_2}{2}\right] \\
&= \frac{1}{2}\left(E[X_1] + E[X_2]\right)] \\
&= \frac{1}{2} \times \frac{2}{\lambda} \\
&= \mu
\end{aligned}$$

Esto implica que $\widehat{\theta_1}$ es insesgado respecto de μ,calculemos su varianza

$$V[\overline{X}] = V\left[\frac{X_1 + X_2}{2}\right]$$

$$= \frac{1}{4}\left(V[X_1] + V[X_2])\right]$$

$$= \frac{1}{4} \times \frac{2}{\lambda^2}$$

$$= \frac{1}{2\lambda^2}$$

Luego $ECM[\widehat{\theta}_1] = V[\widehat{\theta}_1] = \frac{1}{2\lambda^2}$

Vemos ahora para $\widehat{\theta}_2$

$$ECM[\widehat{\theta}_2] = V[\sqrt{X_1 X_2}] + (E[\sqrt{X_1 X_2}] - \mu)^2$$

de donde

$$V[\sqrt{X_1 X_2}] = E[X_1 X_2] - \left(E[\sqrt{X_1}]E[\sqrt{X_2}]\right)^2$$

Calculemos ahora $E[\sqrt{X}]$ con X exponencial de parámetro λ

$$E[\sqrt{X}] = \int_0^\infty x^{1/2}\lambda e^{\lambda x} dx$$

$$= \frac{\frac{1}{2}\Gamma\left(\frac{1}{2}\right)}{\lambda^{1/2}}$$

$$= \left(\frac{\pi}{\lambda}\right)^{1/2} \times \frac{1}{2}$$

Por lo tanto

$$V[\sqrt{X_1 X_2}] = \frac{1}{\lambda} \times \frac{1}{\lambda} - \left(\frac{\pi}{4\lambda}\right)^2$$

$$= \frac{16 - \pi^2}{16\lambda^2}$$

y

$$B^2\left[\sqrt{X_1X_2}\right] = \left(\frac{\pi}{4\lambda} - \frac{1}{\lambda}\right)^2$$

$$= \left(\frac{\pi - 4}{4\lambda}\right)^2$$

De aquí, el Error Cuadrático Medio de $\widehat{\theta_2}$ está dado por

$$ECM[\widehat{\theta_2}] = \frac{16 - \pi^2}{16\lambda^2} + \left(\frac{\pi - 4}{4\lambda}\right)^2$$

$$= \frac{4 - \pi}{2\lambda^2}$$

Como $4 - \pi < 1$ tenemos $ECM[\widehat{\theta_2}] < ECM[\widehat{\theta_1}]$, y de acuerdo a este criterio, $\widehat{\theta_2}$ es preferido a $\widehat{\theta_1}$.

PROBLEMA 6[5]

Sea X_1 y X_2 una muestra aleatoria de tamaño 2 proveniente de una población X con media μ y varianza σ^2.

a) Si disponemos de dos estimadores para $\widehat{\mu_1} = \frac{X_1+X_2}{2}$ y $\widehat{\mu}_2 = \frac{X_1+2X_2}{3}$. ¿Cuál de los dos es el mejor?

b) Para un estimador de la forma $\widehat{\mu} = aX_1 + (1-a)X_2$, con $0 \leq a \leq 1$. Determine el valor de a que conduce al mejor estimador de esta forma.

c) Consideremos el estimador $\widehat{\mu} = \frac{X_1+3X_2}{5}$. ¿Es insesgado? . Si no lo fuera encuentre uno a partir de este estimador.

SOLUCIÓN

a) Veamos el sesgo de los 2 estimadores

- $$E[\widehat{\mu_1}] = E\left[\frac{X_1 + X_2}{2}\right] = \frac{1}{2}(\mu + \mu) = \mu$$

[5] I3 segundo semestre de 2003

- $$E[\widehat{\mu_2}] = E\left[\frac{X_1 + 2X_2}{3}\right] = \frac{1}{3}(\mu + 2\mu) = \mu$$

Los dos estimadores son insesgado, veamos cual de ellos tiene menor varianza

- $$V[\widehat{\mu_1}] = V\left[\frac{X_1 + X_2}{2}\right] = \frac{1}{4}(\sigma^2 + \sigma^2) = \frac{1}{2}\sigma^2$$

- $$V[\widehat{\mu_2}] = V\left[\frac{X_1 + 2X_2}{3}\right] = \frac{1}{9}(\sigma^2 + 4\sigma^2) = \frac{5}{9}\sigma^2$$

Aquí $\widehat{\mu_1}$ tiene menor varianza, y de acuerdo a este criterio, $\widehat{\mu_1}$ es preferido a $\widehat{\mu_2}$.

b) Es insesgado para todo a, en efecto

$$E[\widehat{\mu}] = E[aX_1 + (1-a)X_2]$$

$$= aE[X_1] + (1-a)E[X_2]$$

$$= \mu$$

Veamos su varianza

$$
\begin{aligned}
V[\widehat{\mu}] &= V[aX_1 + (1-a)X_2] \\
&= a^2 V[X_1] + (1-a)^2 V[X_2] \\
&= a^2\sigma^2 + (1-a)^2\sigma^2 \\
&= \sigma^2 \underbrace{(2a^2 - 2a + 1)}_{(*)}
\end{aligned}
$$

Determinemos el valor mínimo que puede tomar $(*)$, lo que implica que es de menor varianza

$$
\begin{aligned}
f(a) &= 2a^2 - 2a + 1 \\
f'(a) &= 4a - 2 &= 0 \\
\Rightarrow \quad a &= \tfrac{1}{2}
\end{aligned}
$$

Como $f''(a) = 4 > 0$, entonces $a = \frac{1}{2}$ es un mínimo.

c)

$$
\begin{aligned}
E[\widehat{\mu}] &= E\left[\frac{1}{5}(X_1 + 3X_2)\right] \\
&= \frac{1}{5}(E[X_1] + 3E[X_2]) \\
&= \frac{4\mu}{5}
\end{aligned}
$$

$\widehat{\mu}$ no es insesgado. Ahora tenemos $E[\widehat{\mu}] = \frac{4\mu}{5}$, esto implica $E\left[\frac{5}{4}\widehat{\mu}\right] = \mu$, luego $\tilde{\mu} = \frac{5}{4}\widehat{\mu} = \frac{1}{4}(X_1 + 3X_2)$ es un estimador insesgado para μ.

PROBLEMA 7[6]

Cierto tipo de componente electrónico tiene una duración Y(en horas)con función de densidad dada por

[6] I3 recuperativa de 2001

$$f(y) = \begin{cases} \frac{1}{\theta^2} y e^{-\frac{y}{\theta}} & y > 0; \theta > 0 \\ 0 & e.o.c. \end{cases}$$

Suponga que tres de tales componentes, al probarlos de manera independiente, presentan duración de 120, 128 y 130 horas.

a) Obtenga el estimador por método de momentos de θ, considerando una m.a. (n)

b) Analice si el estimador encontrado en a) es insesgado. ¿Cuál es la varianza de este estimador?

c) Utilice los valores numéricos que se dan para obtener la estimación de θ.

SOLUCIÓN

a) • Momento muestral : $\overline{x}$

• Momento Población: $E[X] = \frac{\alpha}{\beta} = \frac{2}{1/\theta} = 2\theta$

$$\begin{aligned} \Rightarrow \quad 2\theta &= \overline{x} \\ \Rightarrow \quad \widehat{\theta} &= \tfrac{\overline{x}}{2} \end{aligned}$$

b)

$$E[\widehat{\theta}] = E\left[\frac{\widehat{x}}{2}\right]$$

$$= \frac{1}{2n} \sum_{i=1}^{n} E[x_i]$$

$$= \theta$$

Por lo tanto $\widehat{\theta}$ es insesgado

$$
\begin{aligned}
V[\widehat{\theta}] &= V\left[\frac{\widehat{x}}{2}\right] \\
&= \frac{1}{4n^2} V\left[\sum_{i=1}^{n} x_i\right] \\
&= \frac{1}{4n^2} \sum_{i=1}^{n} V[x_i] \\
&= \frac{1}{4n^2}(2n\theta^2) \\
&= \frac{\theta^2}{2n}
\end{aligned}
$$

c) 120; 128; 130 $\Rightarrow \overline{x} = 126 \therefore \widehat{\theta} = \frac{\overline{x}}{2}$

5.2 Ejercicios Propuestos

1. [7]Se determina la resistencia al corte de cada una de diez soldaduras eléctricas por puntos de prueba, obteniéndose los siguientes datos (lb/pulg^2):

   392 376 401 367 389 362 409 415 358 375

   a) Si se supone que la resistencia al corte está normalmente distribuida, estime el verdadero promedio de resistencia al corte y su desviación estándar con el método de máxima verosimilitud.

   b) Otra vez, suponiendo una distribución normal, estime el valor de resistencia abajo del cual 95% de todas las soldaduras tendrán sus resistencias. (*Sugerencia:* ¿cuál es el 95to percentil en términos de μ y σ? Ahora utilice el principio de invarianza.)

2. Suponga que se tiene una muestra aleatoria de tamaño 2n tomada de una población X, con $E[X] = \mu$ y $Var[X] = \sigma^2$. Sean:

   $$\overline{X}_1 = \frac{1}{2n}\sum_{i=1}^{2n} X_i \quad y \quad \overline{X}_2 = \frac{1}{n}\sum_{i=1}^{n} X_i$$

   dos estimadores de μ. ¿Cuál es el mejor estimador de μ? Explique su elección.

3. Sean $X_1, \ldots, X_7$ una muestra aleatoria de una población que tiene media μ y varianza σ^2. Considere los siguientes estimadores de μ:

   $$\widehat{\mu}_1 = \frac{X_1 + \cdots + X_7}{7} \quad \widehat{\mu}_2 = \frac{2X_1 - X_6 + X_4}{2}$$

   a) ¿Alguno de estos estimadores es insesgado?

   b) ¿Cuál es mejor ? ¿En que sentido es mejor?

4. Suponga que $\widehat{\Theta}_1$ y $\widehat{\Theta}_2$ son estimadores insesgados del parámetro θ. Se sabe que $V(\widehat{\Theta}_1) = 10$ y $V(\widehat{\Theta}_2) = 4$. ¿Cuál es mejor, en que sentido lo es?

5. Calcule la eficiencia relativa de los dos estimadores del ejercicio 2.

6. Calcule la eficiencia relativa de los dos estimadores del ejercicio 3.

7. Suponga que $\widehat{\Theta}_1$ y $\widehat{\Theta}_2$ son estimadores del parámetro θ. Se sabe que $E[\widehat{\Theta}_1] = \theta$, $E[\widehat{\Theta}_2] = \frac{\theta}{2}$, $Var[\widehat{\Theta}_1] = 10$, $Var[\widehat{\Theta}_2] = 4$. ¿Qué estimador es "mejor"? ¿en qué sentido es mejor?

8. Suponga que $\widehat{\Theta}_1$, $\widehat{\Theta}_2$ y $\widehat{\Theta}_3$ son estimadores del parámetro θ. Se sabe que $E[\widehat{\Theta}_1] = \theta$, $E[\widehat{\Theta}_2] = \theta$, $E[\widehat{\Theta}_3] \neq \theta$, $Var[\widehat{\Theta}_1] = 12$, $Var[\widehat{\Theta}_2] = 10$ y $E[\widehat{\Theta}_3 - \theta]^2 = 6$. Haga una comparación de estos tres estimadores. ¿Cuál prefiere? ¿Por qué?

[7]I2 segundo semestre de 2002

9. De una población que tiene media μ y varianza σ^2, se toman tres muestras aleatorias se tamaño $n_1 = 20$, $n_2 = 10$ y $n_3 = 8$. Sean S_1^2, S_2^2 y S_3^2 las varianzas muestrales. Demuestre que $S^2 = \frac{20S_1^2+10S_2^2+8S_3^2}{38}$ es un estimador insesgado de σ^2.

10. a) Demuestre que $\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{n}$ es un estimador sesgado de σ^2.
    b) Determine la magnitud del sesgo del estimador.
    c) ¿Qué sucede con el sesgo a medida que aumenta el tamaño n de la muestra?

11. Sea $X_1, \ldots, X_n$ una muestra aleatoria de tamaño n.
    a) Demuestre que $\overline{X}^2$ es un estimador sesgado de μ^2
    b) Determine la magnitud del sesgo en el estimador
    c) ¿Qué sucede con el sesgo a medida que aumenta el tamaño n de la muestra?

12. Examine la siguiente muestra de observaciones de espesor de pintura de baja viscosidad ("Achieving a Target Value for a Manufacturing Process: A Case Study", *J. of Quality Technology*, 1992, pp. 22-26):

0.83	0.88	0.88	1.04	1.09	1.12	1.29	1.31
1.48	1.49	1.59	1.62	1.65	1.71	1.76	1.83

Suponga que la distribución de espesores de pintura es normal (una gráfica de probabilidad normal respalda esta hipótesis).

a) Calcule un estimado puntual del valor promedio del espesor de pintura y diga qué estimador usó.
b) Calcule un estimado puntual de la mediana de la distribución de espesores de pintura y diga qué estimador usó.
c) Calcule un estimado puntual del valor que separa 10% de los valores más altos de espesores, del restante 90%, y diga qué estimador usó. [*Sugerencia:* exprese lo que trata de estimar en términos de μ y de σ]
d) Estime $P(X < 1.5)$, es decir, la proporción de todos los valores de espesor menores que 1.5.[*Sugerencia:* si conociera los valores de μ y de σ, podría calcular esta probabilidad. Estos valores no están disponibles, pero se pueden estimar.]
e) ¿Cuál es el error estándar estimado del estimador que usó en el inciso b)?

Resp: a) 1.348, $\overline{X}$;b) 1.348, $\overline{X}$; c) 1.781, $\overline{X}$+1.28S ; d)0.6736; e) 0.0905

13. Represente por $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución de Rayleight con pdf

$$f(x;\theta) = \frac{x}{\theta}e^{\frac{-x^2}{2\theta}} \quad x > 0$$

a) Se puede demostrar que $E(X^2) = 2\theta$. Utilice este hecho para construir un estimador insesgado de θ con base en $\sum_i X_i^2$ (y use reglas de valor esperado para demostrar que es insesgado).

b) Estime θ de las siguientes $n = 10$ observaciones sobre esfuerzo vibratorio de una paleta de turbina bajo condiciones específicas:

16.88	10.23	4.59	6.66	13.68
14.23	19.87	9.40	6.51	10.95

Resp: a)$\widehat{\theta} = \frac{\sum X_i^2}{2n}$; b) 74.505

14. Se observan dos sistemas diferentes de computadora durante un total de n semanas. Represente con X_i el número de descomposturas del primer sistema durante la i-ésima semana y suponga que las X_i son independientes y obtenidas de una distribución de Poisson con parámetro λ_1. De forma similar, represente con Y_i el número de descomposturas del segundo sistema durante la i-ésima semana y suponga independencia en cada Y_i de Poisson, con parámetro λ_2. Obtenga las mle de λ_1, λ_2 y $\lambda_1 - \lambda_2$. [Sugerencia: mediante el uso de independencia, escriba la pmf conjunta (verosimilitud) de las X_i y Y_i juntas.]
Resp: $\widehat{\lambda}_1 = \overline{X}, \widehat{\lambda}_2 = \overline{Y}$ el estimado de $(\lambda_1 - \lambda_2)$ es $\overline{X} - \overline{Y}$

15. Se determina la resistencia al corte de cada una de diez soldaduras eléctricas por puntos de prueba, obteniéndose los siguientes datos (lb/pulg^2):

392	376	401	367	389	362	409	415	358	375

a) Si se supone que la resistencia al corte está normalmente distribuida, estime el verdadero promedio de resistencia al corte y su desviación estándar con el método de máxima verosimilitud

b) Otra vez, suponiendo una distribución normal, estime el valor de resistencia abajo del cual 95% de todas las soldaduras tendrán sus resistencias. (*Sugerencia: ¿*cuál es el 95to percentil en términos de μ y σ? Ahora utilice el principio de invarianza.)

Resp: a) 384.4, 18.86 ; b) 415.42

16. Considere una muestra aleatoria $X_1, X_2, \ldots, X_n$, de la fdp exponencial desplazada

$$f(x; \lambda, \theta) = \begin{cases} \lambda e^{-\lambda(x-\theta)}, & x \geq \theta \\ 0 & \text{en otro caso.} \end{cases}$$

a) Obtenga los estimadores de máxima verosimilitud de θ y λ.

b) Si se hacen $n = 10$ observaciones del avance, que resulten en los valores 3.11, 0.64, 2.55, 2.20, 5.44, 3.42, 10.39, 8.93, 17.82 y 1.30, calcule las estimaciones de θ y λ.

Resp: a) $\widehat{\theta} = min(X_i), \widehat{\lambda} = \frac{n}{\sum(X_i - min(X_i))}$;b) 0.64, 0.202

17. Cuando la desviación muestral estándar S está basada en una muestra aleatoria de una distribución normal de población, se puede demostrar que

$$E(S) = \sqrt{\frac{2}{n-1}} \cdot \frac{\Gamma(\frac{n}{2})\sigma}{\Gamma(\frac{n-1}{2})}$$

Utilícela para obtener un estimado insesgado para σ de la forma cS. ¿Cuál es c cuando n=20?
Resp: 1.0132

18. Sea X una variable aleatoria con la siguiente distribución de probabilidad:

$$f(x) = \begin{cases} (\alpha+1)x^{\alpha}, & 0 < x < 1 \\ 0, & \text{en otro caso.} \end{cases}$$

Encuentre el estimador de máxima verosimilitud de α, basado en una muestra de tamaño n.

19. Considere la distribución de Poisson

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$

Encuentre el estimador de máxima verosimilitud de λ, basado en una muestra de tamaño n.

20. Sea $X_1, \ldots, X_n$ una muestra aleatoria de la distribución Geométrica(p)

$$f(x) = (1-p)^{x-1}p \quad x = 0, 1, 2, \ldots$$

a) Encuentre el EMV para p
b) Encuentre el EMV para $p^3 + 2p^2 + 1$

Capítulo 6

Intervalos de Confianza y Test de Hipótesis

6.1 Ejercicios Resueltos

PROBLEMA 1[1]

Se realizó un experimento considerando 64 pacientes varones de similares características que llegan a un servicio de urgencia con fuertes dolores producidos por cálculos renales.
Se les suministró una dosis de 5 ml. De un nuevo fármaco para calcular tales dolores, midiéndose el tiempo transcurrido hasta que el dolor desaparece completamente. Los resultados del experimento entregaron los siguientes resultados:

$\overline{X} = 20$ minutos; $S = 5$minutos. Además 7 pacientes reaccionaron negativamente por la dosis.

a) Mediante un intervalo de confianza del 95%, encuentre los límites que permitan estimar, el tiempo que tarda el medicamento en eliminar el dolor.
(Asuma que el tiempo medio transcurrido hasta que el dolor desaparezca completamente tiene una distribución normal)

b) Estime mediante un intervalo de confianza del 90%, la proporción de pacientes que reaccionaran de manera negativa ante la suministración de la dosis.

c) Si se toma en consideración la información recopila hasta este momento y se desea construir un intervalo con 90% de confianza para la proporción de casos que reacciona negativamente, de tal manera de lograr un error de estimación del 3% como máximo. ¿Cuál es la cantidad mínima de pacientes que debe constituir el grupo experimental?

SOLUCIÓN

De los datos tenemos $n = 64$; $\overline{X} = 20$ min; $\alpha = 5\% = 0.05$; $S = 5$ min; $t_{1-\frac{\alpha}{2};n-1} = t_{0.975;63} = 1.9983$

[1] I3 segundo semestre de 2000

a)

$$\begin{aligned} \mu &\in \overline{X} \mp t_{1-\frac{\alpha}{2};n-1} \times \frac{S}{\sqrt{n}} \\ \Rightarrow \mu &\in 20 \mp 1.9983 \times \frac{5}{\sqrt{64}} \\ \Rightarrow \mu &\in [18.75106, 21.24893] \end{aligned}$$

b) $\widehat{p} = \frac{7}{64}$; $Z_{1-\frac{\alpha}{2}} = Z_{0.95} = 1.645$ $\left(\widehat{p} = \frac{7}{64}, \widehat{q} = 0.89065\right)$

$$\begin{aligned} p &\in \widehat{p} \mp Z_{1-\frac{\alpha}{2}} \times \sqrt{\frac{\widehat{p}\widehat{q}}{n}} \\ \Rightarrow p &\in \frac{7}{64} \mp 1.645 \times \sqrt{\frac{\frac{7}{64} \times \frac{57}{64}}{64}} \\ \Rightarrow p &\in [0.045197, 0.173552] \end{aligned}$$

c)

$$\begin{aligned} \epsilon &= Z_{1-\frac{\alpha}{2}} \times \sqrt{\frac{\widehat{p}\widehat{q}}{n}} \\ \Rightarrow n &= \frac{Z^2 \widehat{p}\widehat{q}}{\epsilon^2} \\ n &= \frac{(1.645)^2 \frac{7}{64} \frac{57}{64}}{(0.03)^2} \\ &= 292.88 \\ &\approx 293 \end{aligned}$$

PROBLEMA 2[2]

Suponga que a partir de una muestra aleatoria de tamañ 25, se ha podido establecer un intervalo de confianza para la media poblacional que va desde 68 a 72 unidades de medida para un $\alpha = 0.01$. Encuentre un intervalo al 95% de confianza para la media poblacional, asuma que la varianza poblacional es conocida.

SOLUCIÓN

Se sabe que $n = 25$, $I.C(\mu) = [68, 72] \Rightarrow \overline{X} = 70$, $\therefore E = 2$

$$\begin{aligned} \Rightarrow 2 &= Z_{1-\frac{0.01}{2}} \frac{\sigma}{\sqrt{25}} \\ \Rightarrow 2 &= Z_{0.995} \frac{\sigma}{5} \\ \Rightarrow \sigma &= \frac{2 \times 5}{2.575} \\ \Rightarrow \sigma &\approx 3.88 \end{aligned}$$

[2]Examen segundo semestre de 2001

Ahora para un intervalo de confianza del 95% para la media poblacional se tiene

$I.C(\mu) \in [70 \mp 1.96\frac{3.88}{5}] \Rightarrow, \mu \in [68.479; 71.520]$

PROBLEMA 3[3]

Una empresa embotelladora lo contrata a Ud. para realizar algunos análisis estadísticos para su línea de producción, debido a varios reclamos sobre la cantidad de líquido en cada botella. Si la máquina embotelladora sigue una distribución normal con media μ y desviación estándar 10c.c.

a) ¿En cuánto debe ser regulado el llenado medio para que sólo el 25% de las botellas tenga menos de 300 c.c.

b) Construya un I.C. de nivel $\alpha = 0.04$ para el llenado medio(μ), si en una muestra de 4 botellas se observa un promedio de 350 c.c. Este intervalo ¿Captura a μ?

c) ¿Cuál debería ser el tamaño mínimo necesario para estimar μ con un error no mayor a 5 c.c. y con una confianza del 90%? Si la confianza sube al 99.9% ¿Qué sucede con el tamaño muestral?

SOLUCIÓN

a)

$$\begin{aligned} P(X < 300) &= 0.25 \\ \Rightarrow \quad \tfrac{300-\mu}{10} &= Z_{0.25} \end{aligned}$$

Se tiene $Z_{0.25} = -Z_{0.75} = -0.7734 \Rightarrow \mu = 300 + 7.7 = 307.7$

b) Los datos nos entregan la siguiente información $\overline{x} = 350$, $Z_{1-\frac{0.04}{2}} = Z_{0.98} = 2.05$, $\sigma = 10$ y $n = 4$

$\therefore \mu \in 350 \mp 2.05 \times \frac{10}{2} \Rightarrow \mu \in (340, 360)$, no captura a μ(obtenido en a).

c) $Z_{1-\frac{\alpha}{2}} = Z_{1-\frac{0.1}{2}} = Z_{0.95} = 1.645$; $\sigma = 10$; $\epsilon = 5$

$$n \geq \left(\frac{1.645 \times 10}{5}\right)^2 = (3.29)^2 \approx 10$$

Si α disminuye o la confianza aumenta, entonces n debe aumentar (o bien, como $Z_{0.95} = 1.645$ y $Z_{1-\frac{0.001}{2}} = Z_{0.9995} \approx 3.0$ y $n \approx 36$)

[3] I3 Recuperativa de 2000

PROBLEMA 4[4]

Se investiga el grado de la dureza Brinell en dos tipos de aleaciones de magnesio, para lo cual se toman muestras aleatorias, resultando para cada aleación las siguientes durezas de Brinell:

Alineación Nro 1	66.3	63.5	64.9	61.8	64.3	64.7	65.1	64.5	68.4	63.2
Alineación Nro 2	71.3	60.4	62.6	63.9	68.8	70.1	64.8	68.9	65.8	66.2

Suponiendo que los distintos tipos de aleación de magnesio de la dureza Brinell se distribuye normal.

a) Calcule un intervalo de confianza para la dureza media de Brinell en aleaciones de magnesio tipo 1, asuma que la varianza poblacional coincide con la varianza muestral(use $\alpha = 10\%$)

b) Encuentre un intervalo de confianza para la proporción de aleaciones de magnesio tipo 1, que tienen una dureza de Brinell inferior a 64.8 con un 95% de nivel de confianza.

c) Determine el tamaño de muestra para estimar la proporción de aleaciones de magnesio tipo 2, con una dureza de Brinell inferior a 64.9, con un 95% de confianza y un error de estimación inferior a 0.20.

SOLUCIÓN

a) De la tabla se obtiene $\overline{X_1} = 64.670$, $\sigma_1 = 1.787$, $n_1 = 10$, $\alpha = 0.1 \Rightarrow Z_{0.95} = 1.645$

$\therefore \mu \in [63.7404; 65.5995]$

b) $\widehat{p} = \frac{6}{10} = 0.6$, $\alpha = 0.05 \Rightarrow Z_{0.975} = 1.96$

$\therefore p \in [0.2964; 0.9036]$

c) $\widehat{p}_2 = 0.4$

$$\begin{array}{rcl} E = Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}_2\widehat{q}_2}{n}} & \leq & 0.20 \\ 1.96\sqrt{\frac{0.4\times0.6}{n}} & \leq & 0.20 \\ n & \geq & \left(\frac{1.96}{0.20}\right)^2 \times 0.6 \times 0.4 \approx 24 \end{array}$$

$\therefore n \geq 24$

PROBLEMA 5[5]

Se piensa que la concentración del ingrediente activo de un detergente líquido para ropa, es afectada por el tipo de catalizador utilizado en el proceso de fabricación. Se sabe que la

[4] Examen segundo semestre de 2001
[5] I3 TAV 2004

desviación estándar de la concentración activa es 3 g/l sin importar el tipo de catalizador utilizado. Se realizan diez observaciones con cada catalizador, y se obtienen los siguientes resultados:

Catalizador 1	57.9	66.2	65.4	65.4	65.2	62.6	67.6	63.7	67.2	71.0
Catalizador 2	66.4	71.7	70.3	69.3	64.8	69.6	68.6	69.4	65.3	68.8

a) Encuentre un intervalo de confianza del 95% para la diferencia entre las medias de las concentraciones activas para los dos catalizadores.

b) ¿Existe alguna evidencia que indique que las concentraciones activas medias dependen del catalizador utilizado?

SOLUCIÓN

a) Los datos nos entrega $\overline{x_1} = 65.22$, $\overline{x_2} = 68.42$, $\sigma_1 = \sigma_2 = 3$, $Z_{\alpha/2} = 1.96$

$$\mu_1 - \mu_2 \quad \in \quad \left[\overline{x_1} - \overline{x_2} \mp Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right]$$

$$\mu_1 - \mu_2 \quad \in \quad [-5.8296, 0.5704]$$

b) Como 0 no esta en el intervalo de confianza, las concentraciones activas medias dependen del catalizador utilizado.

PROBLEMA 6[6]

Una industria dedicada a la fabricación de harina, para llenar los paquetes usa dos máquinas. Se considera que el contenido de harina (kilos) en los paquetes tiene una distribución normal. Para estudiar el contenido de estos paquetes, se toma una muestra aleatoria de cada máquina obteniendo los siguientes resultados:

Máquina A	1.03	1.05	1.08	0.9	1.1	1.2	1.09	1.13	
Máquina B	1.04	1.08	0.9	1	1.06	1.08	1.15	0.92	1.07

a) Se considera que un paquete no cumple con las normas, si su contenido es inferior a un kilo. En base a la muestra total (17 paquetes) y usando un nivel de significancia del 8%, estime la proporción de paquetes que cumplan con la norma. ¿Qué sucede con el tamaño del intervalo anterior si el nivel de significancia es del 5%?

b) La persona encargada de la mantención de la máquina A, sospecha que está no está funcionando correctamente y que existiría una diferencia, respecto del contenido medio de los paquetes llenados por la máquina B. Basándose en la muestra aceptaría Ud. la sospecha del encargado. Use un nivel de confianza del 99%.

[6]Examen segundo semestre de 2003

SOLUCIÓN

a) Sea $\widehat{p} = \frac{14}{17} = 0.82$, $\widehat{q} = 0.18$, $n = 17$

$$p \in \left[\widehat{p} \mp Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\widehat{pq}}{n}}\right] = \left[0.82 \mp 1.755\sqrt{\frac{0.82\times 0.18}{17}}\right]$$

$\therefore p \in [0.656; 0.983]$

Para $\alpha = 0.05$, $Z_{1-\frac{\alpha}{2}} = Z_{0.975} = 1.96$, luego el tamaño aumenta en 0.2

b) De la tabla se obtiene $S_B^2 = 0.006$, $S_A^2 = 0.007$, $\overline{X_A} = 1.072$, $\overline{X_B} = 1.033$, así el $I.C.$

$$I.C\left(\frac{\sigma_B^2}{\sigma_A^2}\right) = \left[\tfrac{0.006}{0.007}F_{7,8,0.005}; \tfrac{0.006}{0.007}F_{7,8,0.995}\right]$$

donde $F_{7,8,0.995} = 7.6942$ y $F_{7,8,0.005} = \frac{1}{F_{8,7,0.995}} = \frac{1}{8.6781} = 0.1152$

$\therefore I.C\left(\frac{\sigma_B^2}{\sigma_A^2}\right) = [0.09874; 6.5950]$

Aquí $1 \in I.C\left(\frac{\sigma_B^2}{\sigma_A^2}\right)$, luego podemos asumir $\sigma_B^2 = \sigma_A^2$

Ahora podemos utilizar $I.C(\mu_A - \mu_B) = \left[(\overline{X_A} - \overline{X_B}) \mp S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}t_{n_1+n_2-2,1-\alpha/2}\right]$

donde $S_p^2 = \frac{7\times 0.007 + 8\times 0.006}{8+9-2} = 0.006466 \Rightarrow 0.0804$

Luego $I.C(\mu_A - \mu_B) = \left[(1.072 - 1.033) \mp 0.0804\sqrt{\frac{1}{8} + \frac{1}{9}}t_{15,0.995}\right]$

$\mu_A - \mu_B \in [-0.076; 0.154] \Rightarrow \mu_A = \mu_B$, con $\alpha = 1\%$

PROBLEMA 7[7]

Supongamos que un fabricante necesita cierta pieza que puede ser proporcionada por dos abastecedores A y B, a un mismo precio. Las piezas de A son defectuosas con probabilidad p_1 y las de B con probabilidad p_2. Supongamos además que de 100 piezas del proveedor A se encontraron 10 piezas defectuosas, mientras que de 150 del proveedor B se encontró 11 defectuosas. ¿Cuál es el provedor con menor proporción de piezas defectuosas.

SOLUCIÓN

De los datos se tiene $\widehat{p_1} = \overline{x} = \frac{10}{100} = 0.01$, $\widehat{p_2} = \overline{y} = \frac{9}{150} = 0.06$, $Z_{0.95} = 1.64$

[7]Examen recuperativo segundo semestre de 2003

Así,

$p_1 - p_2 \in \left[(0.10 - 0.06) \mp 1.64 \left(\frac{0.10\times0.90}{100} + \frac{0.06\times0.94}{150}\right)^{1/2}\right]$

$\therefore\ p_1 - p_2 \in [-0.0186; 0.986]$, como este intervalo contiene al cero, no podemos establecer cual es el proveedor con menor proporción de piezas defectuosas.

PROBLEMA 8[8]

El método usual para tratar la leucemia mieloblástica aguda es tratar al paciente intensamente con quimioterapia en el momento del diagnóstico. Históricamente esto ha producido una tasa de remición del 70%. Estudiando un nuevo método de tratamiento, se utilizaron 50 voluntarios.¿Cuántos de los pacientes deberían haber remitido para que los investigadores pudiesen afirmar con $\alpha = 0.025$, que el nuevo método produce remisiones más altas que el antiguo?

SOLUCIÓN

$p = \frac{70}{100} = 0.7$ histórico

$$\begin{array}{lcccc} H_0 & : & p & = & 0.7 \\ H_1 & : & p & > & 0.7 \end{array}$$

Se quiere rechazar H_0, es decir $T > Z_{1-\alpha}$, donde $Z_{1-\alpha} = Z_{1-0.025} = Z_{0.975} = 1.96$

$$T = \frac{\widehat{p} - p_0}{\sqrt{\frac{p_0 q_0}{50}}} = \frac{\widehat{p} - 0.7}{\sqrt{\frac{0.7\times0.3}{50}}}$$

$\therefore$ para que $T > Z_{1-\alpha} \Rightarrow \frac{\widehat{p}-0.7}{\sqrt{\frac{0.7\times0.3}{50}}} > 1.96 \Rightarrow \widehat{p} > 1.96\sqrt{\frac{0.7\times0.3}{50}} + 0.7 \Rightarrow \widehat{p} > 0.8270$

$\therefore\ n\widehat{p} = 50 \times 0.8270 = 41.35$

Luego deben ser al menos 42 pacientes.

PROBLEMA 9[9]

Veinticuatro animales de laboratorios con deficiencia de vitamina D se dividieron en dos grupos de igual tamaño. El grupo 1 recibió un tratamiento consistente en una dieta que proporcionaba vitamina D. El grupo 2 no fue tratado. al término del período experimental, se hicieron las determinaciones de calcio en el suero sanguíneo de estos animales, obteniéndose los siguientes resultados:

$$\begin{array}{lcclccc} \text{Grupo 1} & \overline{X_1} & = & 11.1mg/100cc & ; & S_1 & = & 2.0mg/100cc \\ \text{Grupo 2} & \overline{X_2} & = & 7.8mg/100cc & ; & S_1 & = & 1.5mg/100cc \end{array}$$

[8] I3 Recuperativa segundo semestre de 2000
[9] Examen segundo semestre de 2000

a) Se espera que el grupo que recibe la dieta que proporciona vitamina D, se observe una cantidad promedio de calcio que supere significativamente los $10.8mg/100cc$. Realice el contraste correspondiente, considerando un nivel de significación del 5%. Concluya.

b) Para el contraste:
H_0 : El contenido medio de calcio en ambos grupos es similar (versus)
H_1 : El contenido medio de calcio del grupo 1 es significativamente mayor al del grupo 2.
¿Cuál sería la conclusión de este experimento con un $\alpha = 0.10$?

SOLUCIÓN

a)
$$\begin{array}{lcl} H_0 & : & \mu = 10.8 \\ H_1 & : & \mu > 10.8 \end{array}$$

$T = \frac{\overline{x}-\mu_0}{S/\sqrt{n}} = \frac{11.1-10.8}{2/\sqrt{12}} = 0.519615$ y $t_{1-\alpha;n-1} = t_{1-0.05;11} = t_{0.95;11} = 1.7959$, con esto se puede ver que $T \not> t_{1-\alpha}$, por lo tanto no podemos rechazar H_0, es decir, no evidencia estadística para pensar que supere significativamente los $10.8mg/100cc$.

b) Nuestra prueba de hipótesis de interés es

$$\begin{array}{lcl} H_0 & : & \mu_1 = \mu_2 \\ H_1 & : & \mu_1 > \mu_2 \end{array}$$

¿Las varianzas poblacionales serán iguales? para responder a esto haremos la siguiente prueba de hipótesis

$$\begin{array}{lcl} H_0 & : & \sigma_1^2 = \sigma_2^2 \\ H_1 & : & \sigma_1^2 \neq \sigma_2^2 \end{array}$$

El estadístico de prueba es $F = \frac{S_1^2}{S_2^2} = \frac{(2.0)^2}{(1.5)^2} = 1.77$ y el valor de tabla es $F_{n_1-1,n_2-1,\alpha/2} = F_{11,11,0.05} = \frac{1}{2.8176} = 0.3549$ o $F_{n_1-1,n_2-1,1-\alpha/2} = F_{11,11,0.95} = 2.871$. Como $F = 1.77 \not> F_{11,11,0.95} = 2.871$ o $F = 1.77 \not< F_{11,11,0.05} = 0.3549$, entonces no hay evidencia estadística de rechazar H_0, luego $\sigma_1^2 = \sigma_2^2$.

Volvamos a nuestra prueba de hipótesis de interés, sabiendo ahora que las varianzas poblaciones son iguales y desconocidas

$T = \frac{\overline{x}_1-\overline{x}_2}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{11.1-7.8}{1.7677\sqrt{\frac{1}{12}+\frac{1}{12}}} = 4.5727$, donde $S_p = \sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}} = \sqrt{\frac{11(2^2+1.5^2)}{22}} = 1.7677$ y $t_{1-\alpha;n_1+n_2-2} = t_{0.9,22} = 1.3212$, y como $T = 4.5727 > t_{0.9,22} = 1.3212$, rechazamos H_0, es decir, el contenido medio de calcio del grupo 1 es significativamente mayor al del grupo 2.

PROBLEMA 10[10]

El presidente de una empresa debe seleccionar un plan, de entre dos que se le presenta, para mejorar la seguridad de sus empleados: el plan A y el plan B. Para ayudarle a tomar una decisión, 9 expertos en seguridad examinan ambos planes y a cada uno se le pide que los clasifiquen en una escala de uno a diez(las clasificaciones altas son para los mejores planes). La compañía adoptará el plan B, que es más caro, sólo si los datos respaldan la evidencia de que los expertos califican mejor el plan B que al plan A. En la tabla se presentan los resultados del estudio. ¿Aportan los datos evidencia de que al nivel $\alpha = 5\%$ las calificaciones del plan B tienden a ser mayores que las del plan A? ¿Para qué valores de α el test es significativo? (Asuma normalidad).

	Juez										
Plan	1	2	3	4	5	6	7	8	9	Prom.	D.E.
A	10	7	8	7	6	7	9	6	7	7.4	1.3
B	9	10	9	8	10	9	8	8	10	9.0	0.8

SOLUCIÓN

Muestras pareadas

$$H_0 \ : \ \mu_d \ \leq \ 0$$

$$H_1 \ : \ \mu_d \ \geq \ 0 \qquad \mu_d = \mu_B - \mu_A$$

$t_c = \frac{\bar{d}}{S_d/\sqrt{n}} = \frac{1.6}{1.8/\sqrt{9}} \approx 2.68$, $t_{0.05};8 = 1.86$ como $t_c > 1.86$ se rechaza H_0

valor-p=$P(t_8 > 2.68) = 1 - P(t_8 \leq 2.68) = 1 - 0.986037 = 0.14$, por lo tanto, $\forall\ \alpha > 1.4\%$ el test se rechaza.

PROBLEMA 11[11]

Un fabricante sostiene que el modelo de auto A, tiene un rendimiento promedio de 13 kilómetros por litro de gasolina. Se selecciona una muestra de 9 de éstos vehículos, y cada uno es conducido con un litro de gasolina en las mismas condiciones. La muestra proporciona una media de 12.34 km/lt, con una desviación estándar de 1.26 km/lt. Nos interesa lo siguiente:

a) Para $\alpha = 0.05$, verificar la afirmación del fabricante.

b) Determinar la probabilidad de cometer error tipo II, si el verdadero valor de μ es de 11 km/lt. De acuerdo a esto, ¿qué se puede decir acerca de la decisión tomada en (a)?

Sugerencia: Utilize el hecho $P(T < 7.07) = 0.9999475$, y $P(T < 2.45) = 0.9800316$

[10] I3 TAV 2004
[11] Control 4 segundo semestre de 2003

c) Si el fabricante sostiene que la desviación estándar poblacional es de 1.20 km/lt, realizar la prueba correspondiente.

d) Supongamos que otro fabricante sostiene que el rendimiento promedio del auto de marca A es mayor a lo indicado por el primer fabricante y además suponga que $\sigma = 1.20km/lt$. Si $\mu = 10$ en la hipótesis alternativa.¿Qué tamaño de muestra se requiere para lograr que las probabilidades de error tipo I y II sean ambas iguales a 0.02?

SOLUCIÓN

a) Supongamos que el rendimiento por litro de gasolina del auto de tipo A es un variable con distribución normal.

La prueba planteada está dada por:

$$H_0 : \mu = 13 \qquad H_1 : \mu \neq 13$$

Para un $\alpha = 0.05$ la región critica es $RC = \{|t_c| < t_{1-\alpha/2}\}$, y el valor observado $t_c = \frac{(12.34-13)3}{1.26} = -1.57$.

Luego como $|t_c| = 1.57 \not> t_{0.975} = 2.31$, no podemos rechazar H_0, es decir la afirmación del fabricante es correcta.

b)

$$\begin{aligned}
\beta &= P(\text{No rechazar} H_0|\mu = 11) \\
&= P(|T| < 2.31|\mu = 11) \\
&= P(-2.31 < \tfrac{3(\bar{x}-13)}{1.26} < 2.31|\mu = 11) \\
&= P(12.0298 < \bar{x} < 13.9702|\mu = 11) \\
&= P(\tfrac{(12.0298-11)3}{1.26} < T < \tfrac{(13.9702-11)3}{1.26}) \\
&= P(2.45 < T < 7.07) \\
&= 0.9999475 - 0.9800316 \\
&= 0.01991583
\end{aligned}$$

Dado que cometer error de tipo II es relativamente baja, para un rendimiento real de 11 km/lt, la decisión de no rechazar H_0 en a) es adecuada

c)

$$H_0 : \sigma^2 = (1.20)^2 \qquad H_1 \sigma^2 \neq (1.20)^2$$

El estadístico de razón de verosimilitud es $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{8(1.26)^2}{1.20^2} = 8.82$

Como $\chi^2 = 8.82 \not< \chi^2_{0.025;8} = 2.18$ y $\chi^2 = 8.82 \not> \chi^2_{0.975;8} = 17.53$ no podemos rechazar H_0, es decir, lo que sostiene el fabricante es correcto.

d) Para el otro fabricante la prueba de hipótesis es:

$$H_0 : \mu \geq 13 \qquad H_1 : \mu < 13$$

Para α se tiene

$$\alpha = 0.02 = P(\bar{X} < c|\mu = 13)$$

$$0.02 = \Phi\left(\frac{c-13}{1.20/\sqrt{n}}\right) \qquad (1)$$

Para β

$$\beta = 0.02 = P(X \bar{\geq} c|\mu = 10)$$

$$0.98 = \Phi\left(\frac{c-10}{1.20/\sqrt{n}}\right) \qquad (2)$$

Luego de (1) y (2) se tiene c=11.5 y $n = 2.6896$, así el tamaño de muestra que se requiere para lograr que las probabilidades de error tipo I y II sean ambas iguales a 0.02 es de $n = 3$.

6.2 Ejercicios Propuestos

1. En determinada empresa de productos químicos , durante un proceso de control de calidad , se encontró que 12 de 100 presentaban defectos

   a) Encuentre un intervalo de confianza del 99% para la proporción de defectuosos en el proceso.

   b) Con un 99% de confianza, ¿cuál es el posible error si la proporción es estimada por 0.12?

2. Supongamos que la longitud de los clavos producidos por una máquina constituye una v.a con distribución normal. Una muestra de 5 clavos proporciona la siguiente información en cuanto a longitud(en pulgadas): 1.14, 1.14, 1.15, 1.12, 1.10.

   a) Construir un intervalo de confianza del 99% para la longitud media de los clavos producidos por esta máquina.

   b) Construir un intervalo de confianza del 90% para la varianza poblacional

3. La probabilidad que una plancha de Zinc fabricada por una máquina sea declarada de "segunda clase", a causa de algún defecto, es p(desconocido).

   a) Determine el EMV de p, basado en los valores observados de una muestra de 1000 planchas fabricadas por esta máquina.

   b) Si en 1000 planchas seleccionadas al azar en un día de porducción se encuentra que 30 son de segunda, determine un intervalo de confianza del 95% para p.

   c) Determine el número de planchas requerida para asegurar con una confianza de 0.95 que el error en la estimación de la proporción de planchas de seguna clase , no sobrepase de 0.02.

4. El banco A seleccionó una muestra al azar de 250 personas de entre sus 10000 clientes con cuenta corriente. Al mismo tiempo y en forma independiente, el banco B seleccionó al azar 200 personas de entre sus 5000 clientes con cuenta corriente. El banco A encontró que 89 personas en esta muestra utilizaban regularmente otros srvicios del banco, mientras que el banco B encontró que 52 personas de la muestra utilizaban otros servicios del banco. Estime la diferencia en la proporción de clientes con cuentas corrientes que regularmente usan otros servicios del banco, en los bancos A y B. Use $\alpha = 0.02$

5. Se utilizan dos máquinas para llenar botellas de plástico con detergente para máquinas lavaplatos. Se sabe que las desviaciones estándar del volumen de llenado son $\sigma_1 = 0.10$ onzas de líquido y $\sigma_2 = 0.15$ onzas de líquido para las dos máquinas, respectivamente. Se toman dos muestras aleatorias de $n_1 = 12$ botellas de la maquina 1 y $n_2 = 10$ botellas de la maquina 2. Los volúmenes promedio de llenado son $\bar{x}_1 = 30.87$ onzas de liquido y $\bar{x}_2 = 30.68$ onzas de liquido.

   a) Construya un intervalo de confianza bilateral del 90% para la diferencia entre las medias del volumen de llenado.

b) Construya un intervalo de confianza del 95% para la diferencia entre las medias del volumen de llenado. Compare el ancho de este intervalo con el ancho del calculado en el inciso a).

6. Para cada una de las siguientes aseveraciones establezca si es una hipótesis estadística legítima y por qué.

   a) $H : \sigma > 100$
   b) $H : \bar{x} = 45$
   c) $H : s \leq .20$
   d) $H : \sigma_1/\sigma_2 < 1$
   e) $H : \bar{X} - \bar{Y} = 5$
   f) $H : \lambda \leq 0.01$, donde λ es el parámetro de una distribución exponencial empleada para un modelo de duración de componentes.

7. Se ha propuesto un nuevo diseño para el sistema de frenos de cierto tipo de automóvil. Si se sabe que para el sistema actual el verdadero promedio de distancia de frenado a 40 millas por hora (mph), bajo condiciones especificadas, es 120 pies. Se propone que el nuevo diseño se ponga en práctica sólo si los datos muestrales indican de manera contundente una reducción en el verdadero promedio de distancia de frenado para el nuevo diseño.

   a) Defina el parámetro de interés y establezca las hipótesis pertinentes.
   b) Suponga que la distancia de frenado para el nuevo sistema está normalmente distribuida con $\sigma = 10$. Represente con $\bar{X}$ el promedio muestral de distancia de frenado para una muestra aleatoria de 36 observaciones. ¿Cuál de las siguientes regiones de rechazo es apropiada: $R_1 = \{\bar{x} : \bar{x} \geq 124.80\}$, $R_2 = \{\bar{x} : \bar{x} \leq 115.20\}$, $R_3 = \{\bar{x} : \bar{x} \geq 125.13 o \bar{x} \leq 114.87\}$?
   c) ¿Cuál es el nivel de significancia más adecuado para la región de la parte b)? ¿Cómo cambiaría la región para obtener una prueba con $\alpha = 0.001$?
   d) ¿Cuál es la probabilidad de que el nuevo diseño no se ponga en práctica cuando su verdadero promedio de distancia de frenado es en realidad 115 pies y la región apropiada de la parte (b) sea utilizada?

8. Antes de convenir en la compra de un pedido grande de hojas de polietileno, para un tipo de cables eléctricos de alta presión, llenos de aceite para submarino, una compañía desea ver evidencia concluyente de que la verdadera desviación estándar de grosor del forro es menor de 0.05 mm. ¿Cuáles hipótesis deben probarse y por qué? En este contexto, ¿Cuáles son los errores tipo I y tipo II?

9. Una muestra aleatoria de 150 donaciones recientes en un banco de sangre revela que 92 eran de sangre tipo A. ¿Sugiere esto que el porcentaje real de donadores tipo A difiere del 40%,el porcentaje de la población con sangre tipo A?. Haga una prueba de hipótesis adecuada,con un nivel de signi.cancia de 0.01. ¿Sería distinta su conclusión si se usara un nivel de significancia de 0.05?.

10. El artículo "An Evaluation of Football Helmets Under Impact Conditions"(*Amer. J. Sport Medicine*, 1984, pp. 233-237) reporta que cuando se sometió a cada casco de fútbol, de una muestra aleatoria de 37 del tipo de suspensión, a cierta prueba de impacto, 24 mostraron daños. Sea p la proporción de todos los cascos de este tipo que muestran daños al probarse de la manera descrita.

    a) Calcule un intervalo de confianza de 99% para p.

    b) ¿Qué tamaño de muestra se requiriría para que el ancho de un intervalo de confianza de 99% fuera 0.10 a lo sumo, independientemente de $\widehat{p}$?

11. Se hicieron las siguientes observaciones de resistencia a la fractura de placas base de 18% de acero maragizado al níquel [" Fracture Testing of Weldments", ASTM Special Publ. No. 381, 1965, pp. 328-356 (en ksi $\sqrt{pulg}$, dadas en orden creciente)]:

69.5	71.9	72.6	73.1	73.3	73.5	75.5	75.7
75.8	76.1	76.2	76.2	77.0	77.9	78.1	79.6
79.7	79.9	80.1	82.2	83.7	93.7		

    Calcule un intervalo de confianza de 99% para la desviación estándar de la distribución de la resistencia a la fractura. ¿Es válido este intervalo, cualquiera que sea la naturaleza de la distribución? Explique.

12. Dos empresas distintas desean establecerse en cierta región y brindar servicios de televisión por cable. Denote por p la proporción de suscriptores potenciales registrados que prefieren la primera empresa sobre la segunda. Considere probar $H_0 : p = 0.5$ contra $H_a : p \neq 0.5$, con base en una muestra aleatoria de 25 individuos. Represente con X el número de suscriptores en la muestra que está a favor de la empresa, y con x el valor observado de X.

    a) ¿Cuál de las siguientes regiones de rechazo es la más adecuada y por qué?
    $R_1 = \{x : x \leq 7 \text{ o } x \geq 18\}$
    $R_2 = \{x : x \leq 8\}$
    $R_3 = \{x : x \geq 17\}$

    b) En el contexto de la situación de este problema, describa cuáles son los errores de tipo I y tipo II.

    c) ¿Cuál es la distribución de probabilidad del estadístico de prueba X cuando H_0 es verdadera? Utilícela para calcular la probabilidad de un error tipo I.

    d) Calcule la probabilidad de un error de tipo II para la región seleccionada cuando p =0.3, de nuevo cuando p =0.4, p =0.6 y p =0.7 .

    e) Mediante el uso de la región seleccionada, ¿qué concluye si 6 de los 25 individuos favoreció a la primera empresa?

Capítulo 7

Test de Homogeneidad, Independencia y Bondad de Ajuste

7.1 Ejercicios Resueltos

PROBLEMA 1[1]

Un investigador desea contrastar ciertas hipótesis que tiene sobre el analfabetismo en comunidades Aymarás. Básicamente, él piensa que en las comunidades rurales el analfabetismo es mayor que en comunidades semi-rurales. Para verificar su hipótesis lleva a cabo encuestas en dos comunidades, en la primera de tipo rural, entrevista aleatoriamente a 200 comuneros. Mientras que, en la comunidad semi-rural se entrevistan a 180 comuneros aleatoriamente. Los resultados obtenidos se presentan en la siguiente tabla.

	Rural	Semi-Rural	Total
Analfabetos	110	115	225
Alfabetos	90	65	155
Total	200	180	380

Plantee las hipótesis correspondiente. ¿Qué diría al investigador basándose en esta información?. Use $\alpha = 5\%$

SOLUCIÓN

H_0 :Existe homogeniedad entre las comunidades rurales y semi-rurales con respecto a su analfabetismo

H_1 :No existe homogeniedad

Los valores esperados son

[1]Examen Recuperativo de 2003

	Rural	Semi-Rural	Total
Analfabetos	118.4	106.6	225
Alfabetos	81.6	73.4	155
Total	200	180	380

Entonces el estadístico del test queda de la siguiente forma

$$\chi_c^2 = \frac{(110-118.4)^2}{118.4} + \frac{(115-106.6)^2}{106.6} + \frac{(90-81.6)^2}{81.6} + \frac{(65-73.4)^2}{73.4} = 3.08$$

El valor de tabla esta dado por $\chi^2_{(I-1)(J-1);1-\alpha} = \chi^2_{1;1,0.95} = 3.84$ y como, $\chi_c^2 = 3.08 \not> \chi^2_{1;1,0.95} = 3.84$, podemos concluir que no existe evidencia suficiente de rechazar H_0, es decir en las comunidades rurales el analfabetismo es similar que en comunidades semi-rurales.

PROBLEMA 2[2]

Se realiza un estudio sobre la asociación(dependencia) entre tipo de hospital y muerte en hospital después de una operación de alto riesgo durante el mes de julio. Fueron seleccionados para su estudio 139 pacientes de operaciones de alto riesgo en hospitales universitarios; otros 528 pacientes fueron escogidos de otro tipos de hospitales. De los pacientes tratados en hospitales universitarios, murieron 32 y de la muestra extraída de los otros hospitales murieron 62.

a) Contrastar la hipótesis nula de no asociación. Para ello construya una tabla adecuada al problema y plantee las hipótesis correspondiente.

b) Si se sabe que a nivel nacional la proporción de personas que mueren en este tipo dee operaciones es del orden del 30%. ¿Hay diferencia significativa en relación con los datos obtenidos con los hospitales universitarios?(Recuerde que la hipótesis alternativa es en relación a lo que la muestra sugiere)

(Use el valor $\alpha = 5\%$ tanto para la parte a)como para la parte b)).

SOLUCIÓN

a) H_0 :No hay asociación; H_1 :Hay asociación

		Tipo de Hospital		
		Universitario	Otros	Total
Condición operación	Muerte	32	62	94
	Vive	107	466	573
	Total	139	528	667

Los valores esperados son

[2] I3 segundo semestre de 2000

		Tipo de Hospital		
		Universitario	Otros	Total
Condición	Muerte	19.59	74.41	94
operación	Vive	119.41	453.59	573
	Total	139	528	667

Entonces el estadístico del test queda de la siguiente forma

$$\chi_c^2 = \frac{(32-19.59)^2}{19.59} + \frac{(62-74.41)^2}{74.41} + \frac{(107-119.41)^2}{119.41} + \frac{(466-453.59)^2}{453.59} = 11.56$$

El valor de tabla esta dado por $\chi^2_{(I-1)(J-1);1-\alpha} = \chi^2_{1;1,0.95} = 3.84$ y como, $\chi^2_c = 11.56 > \chi^2_{1;1,0.95} = 3.84$, podemos concluir que existe evidencia suficiente de rechazar H_0, es decir hay asociación.

b)

$$\begin{array}{rcrcl} H_0 & : & p & = & 0.3 \\ H_1 & : & p & < & 0.3(\text{lo que sugieren los datos}) \end{array}$$

$\widehat{p} = \frac{32}{139}$, $Z_c = \frac{\widehat{p}-p_0}{\sqrt{\frac{p_0q_0}{n}}} = \frac{0.23-0.3}{\sqrt{\frac{0.23\times0.77}{n}}} = -1.96$

y $Z_\alpha = -Z_{1-\alpha} = -Z_{1-0.05} = -Z_{0.95} = -1.645$, y como $Z_c < Z_\alpha \Rightarrow$ rechazo H_0 con $\alpha = 5\%$, es decir hay diferencia significativa en relación con los datos obtenidos con los hospitales universitarios.

PROBLEMA 3[3]

Las especificaciones en la producción de tanques de aire, utilizado en inmersión, requiere que los tanques se llenen a una presión promedio de 600 libras por pulgadas cuadradas(psi) con una desviación estándar de 10 psi. Ahora suponga que Ud. ha sido contratado por una importante fábrica de equipos de inmersión, que produce este tipo de tanques y su primera tarea es verificar si la presión de llenado se distribuye normalmente, como lo especifica la reglamentación.
La muestra aleatoria de 1000 tanques que Ud. obtuvo es la siguiente

Presión de llenado(psi)	Menos de 580	580-590	580-590	600-610	610-620	620 y más
N°de Tanques	20	142	310	370	128	30

¿Qué concluiría Ud. respecto a lo especificado por la reglamentación, con un 5% de nivel de significación?

[3]Examen segundo semestre de 2000

SOLUCIÓN

$X(psi)$	n_i	p_i	np_i
menos de 580	20	0.0228	22.8
580-590	142	0.1359	135.9
590-600	310	0.3413	341.3
600-610	370	0.3413	341.3
610-620	128	0.1359	135.9
620 y más	30	0.0228	22.8
	1000	1.0	1000

$$\begin{aligned} H_0 &: X \sim N(600; 100) \\ H_1 &: X \nsim N(600; 100) \end{aligned}$$

$$\begin{aligned}
p_1 &= P(X < 580) &&= P\left(Z < \tfrac{(580-600)}{10}\right) &&= 0.0228 \\
p_2 &= P(580 < X < 590) &&= P\left(Z < \tfrac{(590-600)}{10}\right) - 0.0228 &&= 0.1587 - 0.028 &&= 0.1359 \\
p_3 &= P(590 < X < 600) &&= P\left(Z < \tfrac{(600-600)}{10}\right) - 0.1587 &&= 0.5 - 0.1587 &&= 0.3413 \\
p_4 &= P(600 < X < 610) &&= P\left(Z < \tfrac{(610-600)}{10}\right) - 0.5 &&= 0.8413 - 0.5 &&= 0.3413 \\
p_5 &= P(610 < X < 620) &&= P\left(Z < \tfrac{(620-600)}{10}\right) - 0.8413 &&= 0.9772 - 0.8413 &&= 0.1359 \\
p_6 &= P(X > 620) &&= 1 - P\left(Z < \tfrac{(620-600)}{10}\right) &&= 1 - 0.9772 &&= 0.0228
\end{aligned}$$

Luego $\chi_c^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} = \frac{(20-22.8)^2}{22.8} + \frac{(142-135.9)^2}{135.9} + \frac{(310-341.3)^2}{341.3} + \frac{(370-341.3)^2}{341.3} + \frac{(128-135.9)^2}{135.9} + \frac{(30-22.8)^2}{22.8} = 8.6344$ y como $\chi_c < \chi^2_{1-\alpha} = 11.071 \Rightarrow$ Aceptar H_0, con un $\alpha = 5\%$

PROBLEMA 4[4]

Se observo la duración en horas de 100 pilas de una determinada marca, obteniéndose los siguientes resultados

Horas	< 20	20-40	40-60	60-80	≥ 80
Frecuencia	5	26	34	22	13

[4] I3 TAV 2004

¿Hay evidencia suficiente para rechazar la hipótesis de que los datos siguen una distribución normal de parámetros $\mu = 50$ y $\sigma = 20$. Use $\alpha = 0.05$

SOLUCIÓN

$$\begin{array}{lcl} H_0 & : & X \sim N(50;20^2) \\ H_1 & : & X \nsim N(50;20^2) \end{array}$$

Horas	<20	20-40	40-60	60-80	≥ 80
0_i Frecuencia	5	26	34	22	13
E_i	6.9	24.2	38.3	24.2	6.7
p_i	0.069	0.242	0.383	0.242	0.067

$\chi^2_c = \sum \frac{(O_i - E_i)^2}{E_i} = 7.26$, $\chi^2_{0.05;4} = 9.49$, como $\chi^2_c > \chi^2$ se rechaza H_0

PROBLEMA 5[5]

Una compañía de productos químicos experimenta con cuatro mezclas distintas para un compuesto de un insecticida. Se toman cuatro muestras aleatorias independientes de 200 insectos y se someten a la acción de uno de los cuatro insecticidas, registrándose el número de insectos muertos. Los resultados son los siguientes:

Resultado	Insecticida 1	Insecticida 2	Insecticida 3	Insecticida 4
Muertos	124	147	141	152
Vivos, pero estériles	76	53	59	48

Lleve a cabo un test de hipótesis que permitan decidir si los insecticidas son diferentes en su eficiencia, use $\alpha = 5\%$

SOLUCIÓN

Prueba de homogeneidad

$\chi^2_c = \sum \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 10.72$, $\chi^2_{0.05} = 7.81$, como $\chi^2_c > \chi$ se rechaza H_0, por lo tanto los insecticidas son diferentes en su eficacia.

[5] I3 TAV 2004

FORMULARIO

	$P(X=x)$	$E(X)$	$V(X)$	$M_X(t)$	$R_X(x)$
$X \sim B(p)$	$p^x(1-p)^{1-x}$	p	pq	$q+pe^t$	$0,1$
$X \sim Bin(n,p)$	$\binom{n}{x}p^x(1-p)^{n-x}$	np	npq	$(q+pe^t)^n$	$0,1,...,n$
$X \sim G(p)$	pq^{x-1}	$\frac{1}{p}$	$\frac{q}{p^2}$	$\frac{pet}{1-qe^t}$, si $qe^t<1$	1,2,...
$X \sim Bineg(r,p)$	$\binom{x-1}{r-1}p^r q^{x-r}$	$\frac{r}{p}$	$\frac{rq}{p}$	$[\frac{pe^t}{1-qe}]^r$, si $qe^t<1$	$r,r+1,...$
$X \sim H(M,N,n)$	$\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$	np	$n\frac{M}{N}(\frac{N-M}{N})(\frac{N-n}{N-1})$		$0,1,...,min(M,n)$
$X \sim P(\mu=\lambda t)$	$\frac{\mu^x e^{-\mu}}{x!}$	μ	μ	$e^{\mu(e^t-1)}$	0,1,...
$X \sim U(a,b)$	$\begin{cases} \frac{1}{b-a}, & a<x<b \\ 0, & \text{e.o.c.} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{tb}-e^{at}}{t(b-a)}$	$\mathbb{R}$
$X \sim E(\lambda)$	$\lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda-t}$,$t \leq \lambda$	$0,1,2,...$
$X \sim N(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2	$e^{\mu t+\frac{\sigma^2 t^2}{2}}$	$\mathbb{R}$
$X \sim Gamma(\alpha,\beta)$	$\frac{x^{(\alpha-1)}e^{\frac{-x}{\beta}}}{\Gamma(\alpha)\beta^\alpha}$	$\alpha\beta$	$\alpha\beta^2$	$(1-\beta t)^{-\alpha}$	$x>0$
$X \sim Erlang(r,\lambda)$	$\frac{\lambda^r x^{r-1}e^{-\lambda x}}{(r-1)!}$	$\frac{r}{\lambda}$	$\frac{r}{\lambda^2}$	$(\frac{\lambda}{\lambda-t})^r$	$x>0$

Capítulo 8

Tablas de distribución

8.1 Distribución t de Student

	Magnitud de α en una cola							
gl	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.0005**
1	1.38	1.96	3.08	6.31	12.71	31.82	63.66	636.58
2	1.06	1.39	1.89	2.92	4.30	6.96	9.92	31.60
3	0.98	1.25	1.64	2.35	3.18	4.54	5.84	12.92
4	0.94	1.19	1.53	2.13	2.78	3.75	4.60	8.61
5	0.92	1.16	1.48	2.02	2.57	3.36	4.03	6.87
6	0.91	1.13	1.44	1.94	2.45	3.14	3.71	5.96
7	0.90	1.12	1.41	1.89	2.36	3.00	3.50	5.41
8	0.89	1.11	1.40	1.86	2.31	2.90	3.36	5.04
9	0.88	1.10	1.38	1.83	2.26	2.82	3.25	4.78
10	0.88	1.09	1.37	1.81	2.23	2.76	3.17	4.59
11	0.88	1.09	1.36	1.80	2.20	2.72	3.11	4.44
12	0.87	1.08	1.36	1.78	2.18	2.68	3.05	4.32
13	0.87	1.08	1.35	1.77	2.16	2.65	3.01	4.22
14	0.87	1.08	1.35	1.76	2.14	2.62	2.98	4.14
15	0.87	1.07	1.34	1.75	2.13	2.60	2.95	4.07
16	0.86	1.07	1.34	1.75	2.12	2.58	2.92	4.01
17	0.86	1.07	1.33	1.74	2.11	2.57	2.90	3.97
18	0.86	1.07	1.33	1.73	2.10	2.55	2.88	3.92
19	0.86	1.07	1.33	1.73	2.09	2.54	2.86	3.88
20	0.86	1.06	1.33	1.72	2.09	2.53	2.85	3.85
21	0.86	1.06	1.32	1.72	2.08	2.52	2.83	3.82
22	0.86	1.06	1.32	1.72	2.07	2.51	2.82	3.79
23	0.86	1.06	1.32	1.71	2.07	2.50	2.81	3.77
24	0.86	1.06	1.32	1.71	2.06	2.49	2.80	3.75
25	0.86	1.06	1.32	1.71	2.06	2.49	2.79	3.73
26	0.86	1.06	1.31	1.71	2.06	2.48	2.78	3.71
27	0.86	1.06	1.31	1.70	2.05	2.47	2.77	3.69
28	0.85	1.06	1.31	1.70	2.05	2.47	2.76	3.67
29	0.85	1.06	1.31	1.70	2.05	2.46	2.76	3.66
30	0.85	1.05	1.31	1.70	2.04	2.46	2.75	3.65
∞	0.84	1.04	1.28	1.64	1.96	2.33	2.58	3.29

8.2 Distribución χ^2

	Proporción del Area hasta $+\infty$						
gl	**0.995**	**0.99**	**0.975**	**0.95**	**0.90**	**0.75**	**0.50**
1	0.00	0.00	0.00	0.00	0.02	0.10	0.45
2	0.01	0.02	0.05	0.10	0.21	0.58	1.39
3	0.07	0.11	0.22	0.35	0.58	1.21	2.37
4	0.21	0.30	0.48	0.71	1.06	1.92	3.36
5	0.41	0.55	0.83	1.15	1.61	2.67	4.35
6	0.68	0.87	1.24	1.64	2.20	3.45	5.35
7	0.99	1.24	1.69	2.17	2.83	4.25	6.35
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34
11	2.60	3.05	3.82	4.57	5.58	7.58	10.34
12	3.07	3.57	4.40	5.23	6.30	8.44	11.34
13	3.57	4.11	5.01	5.89	7.04	9.30	12.34
14	4.07	4.66	5.63	6.57	7.79	10.17	13.34
15	4.60	5.23	6.26	7.26	8.55	11.04	14.34
16	5.14	5.81	6.91	7.96	9.31	11.91	15.34
17	5.70	6.41	7.56	8.67	10.09	12.79	16.34
18	6.26	7.01	8.23	9.39	10.86	13.68	17.34
19	6.84	7.63	8.91	10.12	11.65	14.56	18.34

	Proporción del Area hasta $+\infty$						
gl	**0.25**	**0.10**	**0.05**	**0.03**	**0.01**	**0.005**	**0.001**
1	1.32	2.71	3.84	5.02	6.63	7.88	10.83
2	2.77	4.61	5.99	7.38	9.21	10.60	13.82
3	4.11	6.25	7.81	9.35	11.34	12.84	16.27
4	5.39	7.78	9.49	11.14	13.28	14.86	18.47
5	6.63	9.24	11.07	12.83	15.09	16.75	20.51
6	7.84	10.64	12.59	14.45	16.81	18.55	22.46
7	9.04	12.02	14.07	16.01	18.48	20.28	24.32
8	10.22	13.36	15.51	17.53	20.09	21.95	26.12
9	11.39	14.68	16.92	19.02	21.67	23.59	27.88
10	12.55	15.99	18.31	20.48	23.21	25.19	29.59
11	13.70	17.28	19.68	21.92	24.73	26.76	31.26
12	14.85	18.55	21.03	23.34	26.22	28.30	32.91
13	15.98	19.81	22.36	24.74	27.69	29.82	34.53
14	17.12	21.06	23.68	26.12	29.14	31.32	36.12
15	18.25	22.31	25.00	27.49	30.58	32.80	37.70
16	19.37	23.54	26.30	28.85	32.00	34.27	39.25
17	20.49	24.77	27.59	30.19	33.41	35.72	40.79
18	21.60	25.99	28.87	31.53	34.81	37.16	42.31
19	22.72	27.20	30.14	32.85	36.19	38.58	43.82

8.3 Distribución F ($\alpha = 0.05$)

Gl	**Grados de libertad para el numerador**									
den.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	161	199	216	225	230	234	237	239	241	242
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83

Distribución F
($\alpha = 0.05$) (Continuación)

Gl	Grados de libertad para el numerador								
den.	**12**	**15**	**20**	**24**	**30**	**40**	**60**	**120**	∞
1	244	246	248	249	250	251	252	253	254
2	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5
3	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
30	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

8.4 Distribución Normal

	Segunda cifra decimal en z									
z	0.0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998