

# Contenido

1

## DESCRIPCIÓN DE DATOS CON MEDIDAS NUMÉRICAS 1

- 1.1 Aspectos fundamentales de la ciencia 2
  - Propósitos 2
  - Introducción 2
  - Relaciones entre estadística e investigación 7
  - Medición y estadística 9
  - Inferencia estadística y científica 11
  - Estadística e informe científico 13
- 1.2 Descripción de un conjunto de datos con medidas numéricas 14
- 1.3 Medidas de centro 14
  - Ejercicios 18
- 1.4 Medidas de variabilidad 20
  - Ejercicios 24
- 1.5 Sobre la significación práctica de la medición estándar 25
- 1.6 Una medición del cálculo de  $s$  29
  - Ejercicios 31
- 1.7 Mediciones de posición relativa 34
- 1.8 El resumen de cinco números y la gráfica de caja 38
  - Ejercicios 41
  - Repaso del capítulo 44**
  - Tecnología actual 45**
  - Ejercicios suplementarios 48**

2

## PROBABILIDAD 54

- 2.1 Introducción 55
- 2.2 El papel de la probabilidad en estadística 56
- 2.3 Eventos y el espacio muestral 56
- 2.4 Enfoques o escuelas de la probabilidad 59
- 2.5 Axiomas de probabilidad 60
- 2.6 Particiones 62
- 2.7 Cálculo de probabilidades con el uso de eventos sencillos 63
  - Ejercicios 66
- 2.8 Reglas útiles de conteo 68
  - Ejercicios 73

- 2.9 Relaciones de evento y reglas de probabilidad 74
  - Cálculo de probabilidades para uniones y complementos 76
- 2.10 Independencia, probabilidad condicional y la regla de la multiplicación 79
- 2.11 Probabilidad condicional 81
  - Ejercicios 84
- 2.12 Teorema de Bayes 87
  - Ejercicios 90
- 2.13 ¿Eventos mutuamente excluyentes o independientes? 91
- 2.14 Procesos estocásticos 93
  - Cadenas de Markov 94
    - Representación gráfica 94
    - Representación matricial 96
  - Resumen 107
  - Ejercicios 107
- Ejercicios suplementarios 110**

## 3

**ALGUNAS DISTRIBUCIONES DE PROBABILIDAD IMPORTANTES 116**

- 3.1 Introducción 117
- 3.2 Variables aleatorias discretas y sus distribuciones de probabilidad 117
  - Variables aleatorias 117
  - Distribuciones de probabilidad 117
  - La media y desviación estándar para una variable aleatoria discreta 119
  - Ejercicios 123
- 3.3 La distribución binomial de probabilidad 125
  - Ejercicios 133
- 3.4 Distribuciones de probabilidad para variables aleatorias continuas 136
- 3.5 La distribución de probabilidad de Poisson 139
  - Ejercicios 144
- 3.6 La distribución hipergeométrica de probabilidad 145
  - Ejercicios 147
- 3.7 La distribución normal de probabilidad 148
- 3.8 Áreas tabuladas de la distribución normal de probabilidad 149
  - La variable aleatoria normal estándar 149
  - Cálculo de probabilidades para una variable aleatoria normal general 153
  - Ejercicios 156
- 3.9 La aproximación de la distribución de probabilidad binomial a la normal 158
  - Ejercicios 163
- Repaso del capítulo 164**
- Tecnología actual 166**

**Ejercicios suplementarios** 177

**CASO PRÁCTICO: Un misterio: casos de cáncer cerca de un reactor** 185

**CASO PRÁCTICO: "¿Va a calificar por curva?"** 186

4

## **DISTRIBUCIONES MUESTRALES** 187

4.1 Introducción 188

4.2 Planes muestrales y diseños experimentales 188

Ejercicios 191

4.3 Estadísticas y distribuciones muestrales 192

4.4 El teorema central del límite 195

4.5 La distribución muestral de la media muestral 198

Error estándar 199

Ejercicios 202

4.6 La distribución muestral de la proporción muestral 204

Ejercicios 208

4.7 Una aplicación muestral: control estadístico de procesos (opcional) 209

Una gráfica de control para la media del proceso: la gráfica  $\bar{x}$  209

Una gráfica de control para la proporción de piezas defectuosas: la gráfica  $p$  211

Ejercicios 213

**Repaso del capítulo** 215

**Tecnología actual** 216

**Ejercicios suplementarios** 219

**CASO PRÁCTICO: Muestreo de la Ruleta de Monte Carlo** 222

5

## **ESTIMACIÓN DE MUESTRAS GRANDES** 224

5.1 Dónde hemos estado 225

5.2 A dónde vamos: inferencia estadística 225

5.3 Tipos de estimadores 226

5.4 Estimación puntual 227

Ejercicios 232

5.5 Estimación de intervalo 233

Construcción de un intervalo de confianza 234

Intervalo de confianza de muestra grande para una media poblacional  $\mu$  236

Interpretación del intervalo de confianza 237

Intervalo de confianza de muestra grande para una proporción poblacional  $p$  239

Ejercicios 241

5.6 Estimación de la diferencia entre dos medias poblacionales 242

Ejercicios 245

- 5.7 Estimación de la diferencia entre dos proporciones binomiales 248  
Ejercicios 250
- 5.8 Límites de confianza a una cola 252
- 5.9 Selección del tamaño muestral 253  
Ejercicios 257
- Repaso del capítulo 258**
- Ejercicios suplementarios 259**

## 6 INFERENCIA A PARTIR DE MUESTRAS PEQUEÑAS 262

- 6.1 Introducción 263
- 6.2 Distribución  $t$  de Student 263  
Suposiciones detrás de la distribución  $t$  de Student 266
- 6.3 Inferencias de muestra pequeña respecto a una media poblacional 267  
Ejercicios 271
- 6.4 Inferencias de muestra pequeña para la diferencia entre dos medias poblacionales: muestras aleatorias independientes 274  
Ejercicios 280
- 6.5 Inferencias de muestra pequeña para la diferencia entre dos medias: una prueba de diferencia en pares 283  
Ejercicios 288
- 6.6 Inferencias respecto a la varianza poblacional 290  
Ejercicios 296
- 6.7 Comparación de dos varianzas poblacionales 297  
Ejercicios 303
- 6.8 Repaso de suposiciones de muestra pequeña 305  
**Repaso del capítulo 306**  
**Tecnología actual 307**  
**Ejercicios suplementarios 313**

## 7 ANÁLISIS DE VARIANZA 320

- 7.1 El diseño de un experimento 321
- 7.2 ¿Qué es un análisis de varianza? 322
- 7.3 Las suposiciones para un análisis de varianza 322
- 7.4 El diseño completamente aleatorizado: una clasificación en una dirección 323
- 7.5 El análisis de varianza para un diseño completamente aleatorizado 324  
División de la variación total en un experimento 324  
Prueba de igualdad de las medias de tratamiento 327  
Estimación de diferencias en las medias de tratamiento 329  
Ejercicios 332

- 7.6 Clasificación de medias poblacionales 335
  - Ejercicios 338
- 7.7 Diseño de bloque aleatorizado: una clasificación en dos direcciones 339
- 7.8 El análisis de varianza para un diseño de bloque aleatorizado 340
  - División de la variación total en el experimento 340
  - Prueba de igualdad de las medias de tratamiento y de bloque 343
  - Identificación de diferencias en las medias de tratamiento y de bloque 345
  - Algunos comentarios de precaución en bloqueo 346
  - Ejercicios 347
- 7.9 El experimento factorial  $a \times b$ : una clasificación en dos vías 351
- 7.10 El análisis de varianza para un experimento factorial  $a \times b$  353
  - Ejercicios 357
- 7.11 Repaso de las suposiciones del análisis de varianza 361
  - Gráficas residuales 361
- 7.12 Un breve repaso 363
  - Repaso del capítulo 364**
  - Tecnología actual 365**
  - Ejercicios suplementarios 370**
  - CASO PRÁCTICO: ¡Cómo ahorrar dinero en comestibles! 376**

## REGRESIÓN LINEAL Y CORRELACIÓN 377

- 8.1 Introducción 378
- 8.2 Modelo probabilístico lineal simple 378
- 8.3 El método de mínimos cuadrados 381
- 8.4 Un análisis de varianza para regresión lineal 383
  - Ejercicios 386
- 8.5 Prueba de la utilidad del modelo de regresión lineal 389
  - Inferencias respecto a  $\beta$ , la pendiente de la recta de medias 390
  - El análisis de varianza de la prueba  $F$  393
  - Medir la fuerza de la relación: el coeficiente de determinación 393
  - Interpretación de los resultados de una regresión significativa 394
  - Ejercicios 395
- 8.6 Herramientas de diagnóstico para verificar suposiciones de la regresión 398
  - Términos de error dependientes 398
  - Gráficas residuales 398
  - Ejercicios 399
- 8.7 Estimación y predicción usando la recta ajustada 402
  - Ejercicios 406
- 8.8 Análisis de correlación 408
  - Ejercicios 412

- 8.9 Covarianza 414
- Repaso del capítulo 417**
- Tecnología actual 418**
- Ejercicios suplementarios 421**
- CASO PRÁCTICO: ¿Su automóvil está "Hecho en EUA"? 426**

9

**ANÁLISIS DE REGRESIÓN MÚLTIPLE 428**

- 9.1 Introducción 429
- 9.2 El modelo de regresión múltiple 429
- 9.3 Un análisis de regresión múltiple 430
  - El método de mínimos cuadrados 431
  - El análisis de varianza para regresión múltiple 432
  - Prueba de la utilidad del modelo de regresión 433
  - Interpretación de los resultados de una regresión significativa 434
  - Comprobación de suposiciones de regresión 436
  - Uso del modelo de regresión para estimación y predicción 436
- 9.4 Un modelo de regresión polinomial 437
  - Ejercicios 440
- 9.5 Uso de variables predictoras cuantitativas y cualitativas en un modelo de regresión 444
  - Ejercicios 450
- 9.6 Prueba de conjuntos de coeficientes de regresión 453
- 9.7 Interpretación de gráficas residuales 456
- 9.8 Análisis de regresión por pasos 457
- 9.9 Interpretación errónea de un análisis de regresión 458
  - Causalidad 458
  - Multicolinealidad 458
- 9.10 Pasos a seguir al construir un modelo de regresión múltiple 460
  - Repaso del capítulo 460**
  - Tecnología actual 461**
  - Ejercicios suplementarios 463**

10

**ANÁLISIS DE DATOS CATEGÓRICOS 471**

- 10.1 Introducción 472
  - Estudios evaluativos, un enfoque actual 472
  - Los diferentes objetos de la evaluación 473
  - Estudios evaluativos: procedimientos generales 474
  - Áreas de interés del estudio evaluativo 477
  - Programas susceptibles de evaluación 477
  - Interpretación de los resultados 477

- 10.2 Una descripción del experimento 479
- 10.3 Estadístico  $\chi^2$  cuadrada de Pearson 480
- 10.4 Prueba de probabilidades de celda especificada: la prueba de bondad de ajuste 481
  - Ejercicios 483
- 10.5 Tablas de contingencia: una clasificación de dos vías 485
  - La prueba de independencia  $\chi^2$  cuadrada 485
  - Ejercicios 490
- 10.6 Procedimiento *post hoc* 491
  - Coficiente  $\phi$  ( $\Phi$ ) 493
  - Coficiente de contingencia ( $C$ ) 496
  - Prueba de significancia 498
  - Coficiente  $V$  de Kramer 499
- 10.7 Prueba exacta de Fisher 501
- 10.8 Prueba de McNemar 502
- 10.9 Comparación de varias poblaciones multinomiales: una clasificación de dos vías con totales de fila o columna fijos 504
  - Ejercicios 507
- 10.10 La equivalencia de pruebas estadísticas 509
- 10.11 Otras aplicaciones del estadístico de prueba  $\chi^2$  cuadrada 509
  - Repaso del capítulo 511
  - Tecnología actual 511
  - Ejercicios suplementarios 515

**ESTADÍSTICA NO PARAMÉTRICA 521**

- 11.1 Introducción 522
- 11.2 La prueba de suma de rango de Wilcoxon: muestras aleatorias independientes 522
  - Aproximación normal para la prueba de suma de rango de Wilcoxon 526
  - Ejercicios 529
- 11.3 La prueba del signo para un experimento de dos poblaciones 531
  - Aproximación normal para la prueba del signo 532
  - Ejercicios 534
- 11.4 Una comparación de pruebas estadísticas 535
- 11.5 La prueba de rango con signo de Wilcoxon para un experimento de dos poblaciones 536
  - Aproximación normal para la prueba de rango con signo de Wilcoxon 539
  - Ejercicios 540
- 11.6 La prueba  $H$  de Kruskal-Wallis para diseños completamente aleatorizados 542
  - Ejercicios 546
- 11.7 La prueba  $F_r$  de Friedman para diseños de bloque aleatorizados 548

- 11.8 Prueba de Nemenyi 551
  - Ejercicios 552
- 11.9 Prueba de la mediana 554
- 11.10 Coeficiente de correlación de rango 556
- 11.11 Prueba de significancia de  $r_s$  560
- 11.12 Coeficiente tau ( $\tau$ ) de Kendall 561
- 11.13 Coeficiente de concordancia ( $\omega$ ) de Kendall 564
  - Prueba de significancia de  $\omega$  566
- 11.4 Coeficiente de correlación ( $r_{bp}$ ) biserial de punto 566
  - Prueba de significancia de  $r_{bp}$  569
- 11.15 Prueba de Kappa 570
  - Ejercicios 572
- 11.16 Resumen 575
  - Repaso del capítulo 576**
  - Tecnología actual 577**
  - Ejercicios suplementarios 580**
  - CASO PRÁCTICO: ¿Cómo está su nivel de colesterol? 585**

## TEORÍA DE LA RESPUESTA AL ÍTEM 587

- 12.1 Introducción 588
- 12.2 Teoría clásica de los tests en la psicometría 588
  - Supuestos básicos de la teoría de la puntuación verdadera 589
  - Confiabilidad de un test 589
  - Condiciones de paralelismo 590
  - Características de los ítems en la TCT 590
  - Principales limitaciones de la teoría clásica de los tests 591
- 12.3 ¿Qué ofrece la teoría de la respuesta al ítem? 591
  - Curva característica del ítem (CCI) 592
  - Modelo ideal de Guttman y parámetros de un ítem 593
  - Índice de dificultad 593
  - Discriminación de un ítem 594
  - Parámetro de la pseudoadivinación 594
  - Modelo de ojiva normal 594
  - Reparametrización del modelo de ojiva normal 595
  - Modelo logístico de un parámetro o modelo de Rasch 597
  - Modelo logístico de dos parámetros 599
  - Modelo logístico con tres parámetros 599
- 12.4 Principales supuestos de la TRÍ 600
  - Unidimensionalidad del test 600
  - Indeterminación de la escala de rasgo latente 600

- 12.5 Estimación de parámetros del examinado y los ítems 601
  - Método de estimación de máxima verosimilitud 601
  - Estimación de los parámetros:  $a$  y  $b$  602
- 12.6 Función de información 602
  - Usos de la función de información 603
  - Función de información del test 603
- 12.7 Evaluación de bondad de ajuste del modelo 604
  - Interpretación del índice de bondad de ajuste 604
- 12.8 Modelos politómicos de la teoría de la respuesta al ítem 605
  - Modelos politómicos para categorías ordenadas 606
  - Modelo de respuesta graduada 607
  - Ventajas de los modelos politómicos 609
- 12.9 Resumen 610

## **APÉNDICE A LOS ESCRITOS CIENTÍFICOS 611**

## **APÉNDICE B MATRICES 631**

## **ANEXO 643**

## **FUENTES DE DATOS 684**

## **RESPUESTAS A EJERCICIOS SELECCIONADOS 691**

## **GLOSARIO 701**