

TEXTOS DOCENTS

311

MÉTODOS ESTADÍSTICOS APLICADOS

Pedro Sánchez Algarra (autor y coordinador)
Xavier Baraza Sánchez
Ferran Reverter Comas
Esteban Vegas Lozano

Departament d'Estadística



UNIVERSITAT DE BARCELONA



TEXTOS DOCENTS

311

MÉTODOS ESTADÍSTICOS APLICADOS

Pedro Sánchez Algarra (autor y coordinador)
Xavier Baraza Sánchez
Ferran Reverter Comas
Esteban Vegas Lozano

Departament d'Estadística

Publicacions i Edicions



UNIVERSITAT DE BARCELONA



ÍNDICE

INTRODUCCIÓN.....	15
PRÓLOGO	17
1. VARIABLE ALEATORIA	19
1.1. Concepto de variable aleatoria	19
1.2. Función de distribución de probabilidad de una variable aleatoria	19
1.3. Variables aleatorias discretas	22
1.3.1. Propiedades de la función de densidad de probabilidad (o de cuantía) en el caso discreto	22
1.4. Variables aleatorias absolutamente continuas	24
1.4.1. Definición y propiedades.....	24
1.4.2. Comparación del caso discreto con el absolutamente continuo	26
1.5. Distribución conjunta de dos variables aleatorias	27
1.5.1. Distribución bivariantes discretas	28
1.5.2. Funciones de densidad marginales en el caso discreto	29
1.5.3. Distribuciones bivariantes absolutamente continuas.....	30
1.6. Independencia estocástica de variables aleatorias	31
1.6.1. Independencia de dos variables aleatorias en el caso discreto	31
1.6.2. Independencia de dos variables aleatorias en el caso absolutamente continuo.....	31
1.7. Función de distribución condicional de X en el caso discreto	32
1.8. Función de distribución condicional de Y en el caso discreto	32
1.9. Función de densidad condicional de X en el caso discreto	32
1.10. Función de densidad condicional de Y en el caso discreto.....	33
1.11. Función de densidad condicional de X en el caso absolutamente continuo.....	33
1.12. Función de densidad condicional de Y en el caso absolutamente continuo.....	33
1.13. Esperanza matemática	34
1.13.1. Sentido probabilístico de la esperanza matemática.....	34
1.13.2. Esperanza matemática de una variable aleatoria discreta	34
1.13.3. Esperanza matemática de una variable aleatoria absolutamente continua	35
1.13.4. Significado de la esperanza	35

1.13.5. Propiedades de la esperanza matemática	36
1.13.6. Esperanza matemática condicional	37
1.14. Varianza de una variable aleatoria	38
1.14.1. Varianza de una variable aleatoria discreta	38
1.14.2. Varianza de una variable aleatoria absolutamente continua	39
1.14.3. Propiedades de la varianza.	39
1.15. Teorema de Tchebychev	40
1.16. Momentos de una variable aleatoria X respecto del origen	43
1.16.1. Caso discreto	43
1.16.2. Caso absolutamente continuo	44
1.17. Momentos centrales de una variable aleatoria X	44
1.17.1. Caso discreto	44
1.17.2. Caso absolutamente continuo	44
1.18. Momento de dos variables aleatorias respecto del origen	45
1.18.1. Caso discreto	45
1.18.2. Caso absolutamente continuo	45
1.19. Momento centrales de dos variables aleatorias	45
1.19.1. Caso discreto	45
1.19.2. Caso absolutamente continuo	46
1.20. Covarianza de dos variables aleatorias X e Y	46
1.20.1. Propiedades	46
1.21. Coeficiente de correlación lineal de Pearson	47
1.22. Desigualdad de Schwarz	47
1.23. Propiedades del coeficiente de correlación de Pearson	48
 2. DISTRIBUCIONES DISCRETAS	 51
2.1. Introducción	51
2.2. Distribución de Bernouilli	51
2.2.1. Función de cuantía.	51
2.2.2. Función de distribución	51
2.2.3. Cálculo de la esperanza	52
2.2.4. Cálculo de la varianza	52
2.3. Distribución binominal.	53
2.3.1. Función de cuantía.	54
2.3.2. Función de distribución	55
2.3.3. Propiedad fundamental	56
2.3.4. Cálculo de la esperanza y la varianza	56
2.5. Distribución de Poisson	59
2.5.1. Cálculo de la esperanza	60
2.5.2. Cálculo de la varianza	61
2.5.3. Propiedades de la distribución de Poisson.	63

2.5.4. Aproximación de la distribución binominal por la distribución de Poisson	64
2.6. Distribución hipergeométrica.	66
2.6.1. Cálculo de la función de cuantía	67
2.6.2. Cálculo de la esperanza	69
2.6.3. Cálculo de la la varianza	70
2.7. Distribución geométrica o de Pascal	74
2.7.1. Cálculo de la función de cuantía	74
2.7.2. Cálculo de la esperanza	75
2.7.3. Cálculo de la la varianza	77
2.8. Distribución binominal negativa	77
2.9. Aplicaciones informáticas	79
3. DISTRIBUCIONES ABSOLUTAMENTE CONTINUAS	85
3.1. Distribución uniforme	85
3.1.1. Función de distribución	86
3.1.2. Esperanza matemática	87
3.1.3. Varianza	88
3.2. Distribución normal	89
3.2.1. Definición. Función de densidad.	90
3.2.2. Propiedades	92
3.2.3. Función de distribución	94
3.2.4. Distribución normal reducida o estandarizada	94
3.2.5. Áreas bajo la curva normal	97
3.2.6. Tablas de la normal reducida	98
3.2.7. La distribución normal como límite de la distribución binomial	102
3.3. Distribuciones Gamma, Exponencial y Ji-Cuadrado	104
3.3.1. Función Gamma de Euler	104
3.3.2. Distribución Gamma	105
3.3.3. Distribución exponencial	105
3.3.4. Distribución ji-cuadrado de Pearson (χ^2)	107
3.3.5. Esperanza y varianza de la distribución Gamma.	109
3.3.6. Esperanza y varianza de la distribución Exponencial	109
3.3.7. Esperanza y varianza de la distribución ji-cuadrado	110
3.4. Distribución t de Student	110
3.5. Distribución F de Fisher-Snedecor.	112
3.6. Ejemplo con ordenador	115
4. DISTRIBUCIONES DE LOS ESTADÍSTICOS EN EL MUESTREO	123
4.1. Introducción	123
4.2. Muestreo aleatorio simple	123
4.3. Estadístico	125

4.4. Estadístico suma muestral	126
4.5. Estadístico media muestral.	127
4.6. Distribución muestral del estadístico suma muestral a partir de una población normal . .	128
4.7. Distribución muestral del estadístico media muestral a partir de una población normal. .	129
4.7.1. El tamaño de la muestra	129
4.8. Distribución muestral del estadístico suma muestral a partir de una población no normal	130
4.9. Distribución muestral del estadístico media muestral a partir de una población no normal	130
4.10. Distribución muestral de una proporción.	131
4.11. Distribución muestral de la diferencia de dos proporciones	133
4.12. Distribución muestral del estadístico varianza muestral en una población normal	134
4.13. Teorema de Fisher en una población normal	136
4.14. Teorema de Fisher en dos poblaciones normales independientes	138
4.15. Distribución del estadístico cociente de varianzas muestrales corregidas en dos poblaciones normales independientes con igualdad de varianzas	140
 5. ESTIMACIÓN PUNTUAL DE PARÁMETROS Y MÉTODOS DE ESTIMACIÓN	 143
5.1. Introducción	143
5.2. Estimador	143
5.3. Error cuadrático medio del estimador	144
5.4. Propiedades deseables de los estimadores	146
5.4.1. Estimadores centrados o insesgados	146
5.4.2. Cota de Cramér-Rao	149
5.4.3. Estimador de mínima varianza	150
5.4.4. Estimador eficiente	150
5.4.5. Estimador consistente	154
5.4.6. Estimador suficiente	156
5.5. Revisión detallada de los conceptos de verosimilitud e información	160
5.6. Métodos de estimación puntual	166
5.6.1. Método de la máxima verosimilitud	166
5.6.2. Método de los momentos.	171
 6. CONTRASTE DE HIPÓTESIS.	 177
6.1. Introducción	177
6.2. Regiones críticas	179
6.3. Región crítica óptima o prepotente	181
6.3.1. Lema de Neyman-Pearson.	199
6.4. Test de Razón de Verosimilitud	198
6.4.1. Propiedades	199

7. ESTIMACIÓN POR INTERVALOS DE CONFIANZA.	225
7.1. Introducción.	225
7.2. Intervalo de confianza.	225
7.3. Métodos de construcción de intervalos de confianza.	226
7.3.1. Método pivotal.	226
7.3.2. Método general de Neyman.	228
7.4. Intervalos de confianza para una muestra en una población normal.	234
7.4.1. Intervalos de confianza para la media de una población normal, con desviación típica σ conocida.	234
7.4.2. Intervalo de confianza para la media de una población normal con desviación típica σ desconocida y muestra de tamaño pequeño.	236
7.4.3. Intervalo de confianza para la media de una población “no normal” y muestras de tamaño grande.	237
7.4.4. Determinación del tamaño muestral.	238
7.4.5. Intervalo de confianza para la varianza de una población normal con μ desconocida.	239
7.4.6. Intervalo de confianza para la varianza de una población normal con μ conocida.	240
7.4.7. Región de confianza simultánea para (μ, σ^2) en una población normal.	241
7.4.8. Intervalo de confianza de una proporción.	242
7.4.9. Determinación del tamaño muestral para estimar la proporción p de una población.	244
7.5. Intervalos de confianza para dos muestras en dos poblaciones normales.	245
7.5.1. Intervalo de confianza para la diferencia de medias de dos poblaciones normales independientes, con varianzas conocidas.	245
7.5.2. Intervalo de confianza para la diferencia de medias de dos poblaciones independientes no necesariamente normales, con varianzas σ_1^2 y σ_2^2 desconocidas y muestras de tamaño grande.	246
7.5.3. Intervalo de confianza para la diferencia de medias de dos poblaciones normales independientes con varianzas iguales y desconocidas.	247
7.5.4. Intervalo de confianza para la diferencia de medias de dos poblaciones normales independientes con varianzas distintas y desconocidas y muestras de tamaño pequeño.	249
7.5.5. Intervalo de confianza para la diferencia de medias en poblaciones normales: Datos apareados.	250
7.5.6. Intervalo de confianza del cociente de varianzas en poblaciones normales independientes.	251
8. CONTRASTES DE HIPÓTESIS EN EL SUPUESTO DE NORMALIDAD.	255
8.1. Introducción.	255
8.2. Metodología.	255

8.3. El p-valor de un contraste	255
8.4. Contrastes para la media de una población normal	256
8.4.1. Test de dos colas con varianza conocida	256
8.4.2. Test de una cola con varianza conocida	257
8.4.3. Test de dos colas con varianza desconocida	258
8.5. Contrastes para la varianza de una población normal	259
8.6. Contrastes de una proporción.	261
8.7. Test F de comparación de varianzas en dos poblaciones normales independientes.	262
8.8. Contrastes de dos distribuciones normales independientes	264
8.8.1. Contraste de medias con varianzas conocidas	264
8.8.2. Contraste de la diferencia de medias en poblaciones normales independientes con varianzas desconocidas, pero supuestamente iguales: Muestras independientes	265
8.8.3. Contraste de la diferencia de medias en poblaciones normales independientes con varianzas desconocidas y distintas: Muestras independientes	268
8.9. Contraste de igualdad de medias para datos apareados	271
8.10. Comparación de proporciones en muestras apareadas. Variables-respuesta dicotómicas (o dicotomizadas). Test de McNemar	275
8.11. Contraste de la diferencia de proporciones	284
8.12. Ejemplo con ordenador	283
 9. REGRESIÓN LINEAL SIMPLE.	 293
9.1. Introducción	293
9.2. Modelo de regresión lineal simple	294
9.3. Estimadores mínimo-cuadráticos de los coeficientes de la regresión	295
9.3.1. Resumen	297
9.3.2. Otras expresiones de la suma de cuadrados del error	297
9.4. Propiedades de los estimadores β_0 y β_1	298
9.5. Estimación de σ^2	301
9.6. Fórmula fundamental de descomposición de la variabilidad en la regresión lineal y su relación con el coeficiente de correlación de Pearson.	302
9.7. Distribución de los estimadores β_0 y β_1 de los coeficientes de regresión en el muestreo	308
9.8. Distribución del estadístico SSE/σ^2 en el muestreo	309
9.9. Inferencias estadísticas acerca de los parámetros de la regresión	311
9.9.1. Inferencias relacionadas con β_1	312
9.9.2. Inferencias relacionadas con β_0	314
9.9.3. Inferencias relacionadas con $\beta_0 + \beta_1 X$ en $X = x_0$	315
9.9.4. Intervalo de predicción de una respuesta futura en $X = x_0$	317
9.10. Ejemplo con ordenador	319
 TABLAS	 327

PRÓLOGO

Es un hecho incuestionable que los métodos estadísticos constituyen una herramienta imprescindible para aquellos profesionales e investigadores que han de manejar datos y obtener conclusiones de tales informaciones. Por este motivo, la Estadística está pisando con fuerza en ámbitos del conocimiento sumamente diversificados (Ingeniería, Ciencias Experimentales, Medicina, Psicología, Económicas, Sociología, etc.), y para todos ellos se ofrecen conceptos de desarrollo teórico de fácil aplicabilidad.

Con esta obra se facilita al lector un material completo y consistente –teoría, ejercicios y problemas- que contribuya al afianzamiento de sus conocimientos de Estadística, y le aporte un motivo de reflexión sobre los diversos conceptos utilizados y sus aplicaciones. El amplio corpus de conocimiento estadístico ha obligado a un disciplinado esfuerzo de contención para evitar una extensión desmesurada de la obra, dado el desplegamiento de contenidos que ofrece esta disciplina en la actualidad.

El contenido de la obra se estructura en nueve capítulos. En los tres primeros se tratan las variables aleatorias y las diferentes distribuciones discretas y absolutamente continuas, y, son básicos para comprender el cuarto, relativo a la distribución de los estadísticos en el muestreo. En los capítulos cinco y seis se trata el tema de la estimación puntual de parámetros, los diferentes métodos de estimación y la teoría sobre el contraste de hipótesis. Los capítulos siete y ocho se refieren a la estimación por intervalos de confianza y al contraste de hipótesis paramétrico. Finalmente, el capítulo nueve se dedica a la regresión lineal simple.

Diversos capítulos del texto están ilustrados con una aplicación práctica, utilizando el paquete estadístico *Statgraphics Plus*, versión 5.1. Entendemos que esta vertiente relativa a aplicaciones resulta de suma utilidad para quienes pretendan adquirir una sólida formación estadística, y les ofrece un apoyo indudable desde este segundo flanco, sin duda relevante.

Esta obra, fruto de la experiencia docente en diversas titulaciones en las cuales la Estadística tiene un papel relevante, se ha concebido con la idea de que pueda contribuir a los estudiantes de estas titulaciones a adquirir los conocimientos básicos de la disciplina, y teniendo especialmente en cuenta la realidad coyuntural que nos envuelve, en un momento de discusión e implantación de nuevos planes de estudio, que suponen una reducción del tiempo dedicado a la materia, pero no de los contenidos.

1. VARIABLE ALEATORIA

1.1 Concepto de variable aleatoria

Sea un espacio de probabilidades $(\Omega, \mathcal{A}, P)$, donde $\mathcal{A}$ es una álgebra de sucesos. Si el álgebra $\mathcal{A}$ son las “partes de Ω ” ($\mathcal{A} = \wp(\Omega)$), una variable aleatoria X es toda aplicación del espacio muestral Ω en la recta real $\mathbb{R}$.

$$\begin{array}{ccc} X: \Omega & \longrightarrow & \mathbb{R} \\ \in & & \in \\ \omega & \longrightarrow & X(\omega) \end{array}$$

siendo $X(\omega)$ el valor de la variable aleatoria.

Obsérvese que X aplica a cada suceso elemental en un punto de la recta real y que con un mismo Ω se pueden definir infinitas variables aleatorias distintas.

1.2 Función de distribución de probabilidad de una variable aleatoria

$$F(x) = P[X \leq x] \quad \forall x \in \mathbb{R}$$

Obsérvese que esta función solamente existe para variables aleatorias porque si X es una variable aleatoria, el conjunto

$$[X \leq x] = \{\omega \in \Omega \mid X(\omega) \leq x\} \subset \Omega$$

es un suceso.

La función de distribución cumple las siguientes propiedades:

$$1) \quad 0 \leq F(x) \leq 1 \quad \forall x \in \mathbb{R}$$

La demostración es trivial, pues por definición, $F(x)$ es una probabilidad y la misma está comprendida entre 0 y 1.

$$2) \lim_{x \rightarrow -\infty} F(x) = 0; \quad \lim_{x \rightarrow +\infty} F(x) = 1$$

En efecto, se puede demostrar:

$$\begin{aligned} \lim_{x \rightarrow -\infty} F(x) &= \lim_{x \rightarrow -\infty} P[X \leq x] = \lim_{x \rightarrow -\infty} P\{\omega \in \Omega \mid X(\omega) \leq x\} = \\ &= P\left[\lim_{x \rightarrow -\infty} \{\omega \in \Omega \mid X(\omega) \leq x\}\right] = P\{\omega \in \Omega \mid X(\omega) \leq -\infty\} = \\ &= P(\emptyset) = 0 \end{aligned}$$

$$\begin{aligned} \lim_{x \rightarrow +\infty} F(x) &= \lim_{x \rightarrow +\infty} P[X \leq x] = \lim_{x \rightarrow +\infty} P\{\omega \in \Omega \mid X(\omega) \leq x\} = \\ &= P\left[\lim_{x \rightarrow +\infty} \{\omega \in \Omega \mid X(\omega) \leq x\}\right] = P\{\omega \in \Omega \mid X(\omega) \leq +\infty\} = \\ &= P(\Omega) = 1 \end{aligned}$$

3) $F(x)$ es no decreciente. Esto significa que

$$si \quad x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2) \quad \forall x_1, x_2 \in \mathbb{R}$$

En efecto:

$$x_1 < x_2 \Rightarrow \{\omega \in \Omega \mid X(\omega) \leq x_1\} \subset \{\omega \in \Omega \mid X(\omega) \leq x_2\}$$

y teniendo en cuenta que

$$si \quad A \subseteq B \quad entonces \quad P(A) \leq P(B)$$

resulta

$$P[X \leq x_1] \leq P[X \leq x_2] \quad \text{o lo que es lo mismo} \quad F(x_1) \leq F(x_2)$$

$$4) P[a < X \leq b] = F(b) - F(a)$$

En efecto:

$$[X \leq b] = [X \leq a] + [a < X \leq b]$$

Aplicando a los dos miembros de la igualdad la probabilidad, se tiene:

$$P[X \leq b] = P\{[X \leq a] + [a < X \leq b]\}$$

Por ser los sucesos $[X \leq a]$ y $[a < X \leq b]$ mutuamente excluyentes

$$P[X \leq b] = P[X \leq a] + P[a < X \leq b]$$

Despejando $P[a < X \leq b]$ resulta:

$$P[a < X \leq b] = P[X \leq b] - P[X \leq a]$$

y teniendo presente la definición de función de distribución queda:

$$P[a < X \leq b] = F(b) - F(a)$$

que es lo que se quería comprobar.

5) $F(x)$ es continua a la derecha, es decir,

$$\lim_{\substack{x \rightarrow a \\ x > a}} F(x) = F(a) \quad \forall a \in \mathbb{R}$$

Se ha de demostrar que

$$\lim_{\substack{x \rightarrow a \\ x > a}} F(x) - F(a) = 0 \quad \forall a \in \mathbb{R}$$

En efecto:

$$\begin{aligned} \lim_{\substack{x \rightarrow a \\ x > a}} F(x) - F(a) &= \lim_{\substack{x \rightarrow a \\ x > a}} F(x) - \lim_{\substack{x \rightarrow a \\ x > a}} F(a) \\ \lim_{\substack{x \rightarrow a \\ x > a}} F(x) - F(a) &= \lim_{\substack{x \rightarrow a \\ x > a}} [F(x) - F(a)] \end{aligned}$$

Teniendo presente la propiedad (4) resulta:

$$\begin{aligned} \lim_{\substack{x \rightarrow a \\ x > a}} F(x) - F(a) &= \lim_{\substack{x \rightarrow a \\ x > a}} P[a < X \leq x] \\ \lim_{\substack{x \rightarrow a \\ x > a}} F(x) - F(a) &= P \left\{ \lim_{\substack{x \rightarrow a \\ x > a}} [a < X \leq x] \right\} \\ \lim_{\substack{x \rightarrow a \\ x > a}} F(x) - F(a) &= P[a < X \leq a] \\ \lim_{\substack{x \rightarrow a \\ x > a}} F(x) - F(a) &= P\{\omega \in \Omega \mid a < X(\omega) \leq a\} \\ \lim_{\substack{x \rightarrow a \\ x > a}} F(x) - F(a) &= P(a, a] \\ \lim_{\substack{x \rightarrow a \\ x > a}} F(x) - F(a) &= P(\emptyset) = 0 \end{aligned}$$

que es lo que se quería comprobar.

Interesa estudiar dos casos concretos, que son el *caso discreto* y el *caso absolutamente continuo*.

1.3 Variables aleatorias discretas

Se dice que una variable aleatoria X sigue una distribución discreta, si el conjunto de valores que toma con probabilidad no nula, es finito o infinito numerable. Esto quiere decir que existe una sucesión de números reales $x_1, x_2, \dots, x_n, \dots$, tal que:

$$\begin{aligned} P[X = x_i] &= p_i \neq 0 \\ P[X \neq x_i] &= 0 \end{aligned} \quad i = 1, 2, \dots, n, \dots$$

En el caso discreto más que la función de distribución se utiliza la función de densidad de probabilidad que en este caso se define así:

$$\begin{aligned} f(x_i) &= P[X = x_i] = p_i \neq 0 \\ f(x) &= 0 \quad \text{si } x \neq x_i \end{aligned} \quad i = 1, 2, \dots, n, \dots$$

Son ejemplos de variables discretas, el número de accidentes por año en una planta química, el número de piezas defectuosas producidas por una máquina, etc.

1.3.1 Propiedades de la función de densidad de probabilidad (o de cuantía) en el caso discreto

1) $0 \leq f(x) \leq 1 \quad \forall x \in \mathbb{R}$

La demostración es trivial, por ser la función de densidad en el caso discreto una probabilidad.

2) $\sum_{i=1}^{n(o\infty)} f(x_i) = 1$

3) $F(x) = P[X \leq x] = \sum_{x_i \leq x} f(x_i)$

$F(x)$ es la función de distribución de la variable aleatoria X o probabilidad acumulada.

Ejemplo 1.1

Supóngase que en una bolsa hay tres monedas de un €, dos monedas de 50 céntimos de € y una moneda de 10 céntimos de €. Además, que todas las monedas son del mismo tamaño y que no se pueden diferenciar con el tacto. Si se escogen dos monedas, primero una y después otra, sin devolver la primera a la bolsa, y consideramos la variable aleatoria X correspondiente a la ganancia del juego, determinar:

- La función de densidad de probabilidad.
- La función de distribución.
- $P(X = 1)$
- $P(X < 1,5)$
- $P(0,60 < X < 1,5)$

Solución:

- a) Podemos esquematizar el resultado del juego mediante la siguiente tabla

Primera moneda	Segunda moneda	Ganancia	Probabilidad
1€	1€	2€	$3/6 \cdot 2/5 = 1/5$
1€	50cens	1,5€	$3/6 \cdot 2/5 = 1/5$
1€	10cens	1,1€	$3/6 \cdot 1/5 = 1/10$
50cens	1€	1,5€	$2/6 \cdot 3/5 = 1/5$
50cens	50cens	1€	$2/6 \cdot 1/5 = 1/15$
50cens	10cens	60cens	$2/6 \cdot 1/5 = 1/15$
10cens	1€	1,1€	$1/6 \cdot 3/5 = 1/10$
10cens	50cens	60cens	$1/6 \cdot 2/5 = 1/15$

Se trata de una variable discreta con función de densidad

X	$f(x)=P(X=x)$
2	$1/5$
1,5	$2/5$
1,1	$2/10$
1	$1/15$
0,60	$2/15$

- b) Para hallar la función de distribución de X tengamos en cuenta su definición

$$F(x) = P(X \leq x).$$

Entonces

$$F(x) = \begin{cases} 0 & x < 0,60 \\ 2/15 & 0,60 \leq x < 1 \\ 3/15 & 1 \leq x < 1,1 \\ 6/15 & 1,1 \leq x < 1,5 \\ 12/15 & 1,5 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

- c) Observando la función de cuantía del apartado a) tenemos que

$$P(X=1) = 1/15$$

- d) Observando la función de distribución en el apartado b) tenemos que

$$P(X < 1,5) = P(X \leq 1,1) = 6/15$$

- e) Teniendo en cuenta que se trata de una variable aleatoria discreta, la probabilidad pedida viene dada por

$$P(0,6 < X < 1,5) = P(X = 1) + P(X = 1,1) = 4/15$$

1.4 Variables aleatorias absolutamente continuas

Cuando X es una variable aleatoria absolutamente continua, toma una infinidad (no numerable) de valores, por lo que no resulta apropiado hallar la probabilidad de cada valor de la variable, es de esperar que cada uno de ellos tenga probabilidad nula. Por tanto, se admitirá que para cualquier número real a se verifica $P(X = a) = 0$, y sólo son susceptibles de tener probabilidad no nula los intervalos.

Son ejemplos de variables aleatorias absolutamente continuas, la conversión de una reacción química, el espesor de una lámina, el tiempo de vida de un instrumento, etc.

1.4.1 Definición y propiedades

Se dice que una variable aleatoria X es absolutamente continua si existe una función $f(x)$ denominada función de densidad tal que su función de distribución $F(x)$ se puede escribir así:

$$F(x) = \int_{-\infty}^x f(t) dt \quad \forall x \in \mathbb{R}$$

La función $f(x)$ verifica las siguientes propiedades:

1) $f(x) \geq 0$

2) $\int_{-\infty}^{+\infty} f(x) dx = 1$

En efecto, ya que,

$$\int_{-\infty}^{+\infty} f(x) dx = \lim_{x \rightarrow +\infty} \int_{-\infty}^x f(t) dt = \lim_{x \rightarrow +\infty} F(x) = 1$$

3) $F'(x) = f(x)$, cuando $f(x)$ sea continua en x .

4) $P(x < a \leq b) = \int_a^b f(x) dx$

En efecto,

$$P(a < X \leq b) = F(b) - F(a) = \int_{-\infty}^b f(x) dx - \int_{-\infty}^a f(x) dx$$

pero siendo $a < b$,

$$\int_{-\infty}^b f(x)dx = \int_{-\infty}^a f(x)dx + \int_a^b f(x)dx$$

y entonces resulta:

$$P(a < X \leq b) = \int_{-\infty}^a f(x)dx + \int_a^b f(x)dx - \int_{-\infty}^a f(x)dx = \int_a^b f(x)dx$$

La propiedad (1) expresa que la curva representativa de la función de densidad $f(x)$ se encuentra por encima del eje de abscisas. La propiedad (2) expresa que el área total bajo la curva de la función de densidad vale la unidad.

Ejemplo 1.2

Sea X una variable absolutamente continua de función de densidad de probabilidad:

$$f_X(x) = \begin{cases} k \cdot (1 + x^2) & \text{si } x \in (0, 3) \\ 0 & \text{si } x \notin (0, 3) \end{cases}$$

Se pide:

- Hallar la constante k y la función de distribución de probabilidad.
- Probabilidad de que X esté comprendido entre 1 y 2.
- Probabilidad de que X sea menor que 1.
- Sabiendo que X es mayor que 1, probabilidad de que sea menor que 2.

Solución:

a) Se verifica:

$$\int_{-\infty}^0 f(x)dx + \int_0^3 f(x)dx + \int_3^{+\infty} f(x)dx = 1$$

Al ser $f(x) = 0$ si $x \notin (0, 3)$, las integrales primera y tercera son nulas y entonces queda:

$$\int_0^3 f(x)dx = 1 \quad ; \quad \int_0^3 k(1 + x^2)dx = 1 \quad ; \quad k \left[x + \frac{x^3}{3} \right]_0^3 = 1$$

$$k \cdot 12 = 1 \Rightarrow k = \frac{1}{12}$$

La función de distribución asociada a la función de densidad de probabilidad es:

$$F(x) = \int_{-\infty}^x f(t) dt = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{1}{12} \left(x + \frac{x^3}{3} \right) & \text{si } 0 < x < 3 \\ 1 & \text{si } x \geq 3 \end{cases}$$

b)

$$P(1 < X < 2) = \int_1^2 \frac{1}{12} (1 + x^2) dx = \frac{1}{12} \left[x + \frac{x^3}{3} \right]_1^2 = \frac{1}{12} \left(2 + \frac{8}{3} - 1 - \frac{1}{3} \right) = \frac{5}{18}$$

c)

$$P(X < 1) = P(X \leq 1) = F(1) = \frac{1}{12} \left(1 + \frac{1}{3} \right) = \frac{1}{9}$$

d) Es una probabilidad condicionada:

$$P([X < 2] | [X > 1]) = \frac{P([X > 1] \cap [X < 2])}{P(X > 1)} = \frac{P(1 < X < 2)}{1 - P(X \leq 1)} = \frac{5/18}{8/9} = \frac{5}{16}$$

1.4.2 Comparación del caso discreto con el absolutamente continuo

En el caso discreto, $f(x) = P(X = x)$. En el caso absolutamente continuo, $f(x) \neq P(X = x)$.

Es decir, en el discreto, la función de densidad representa una probabilidad y, por tanto, no puede valer más de 1. En el absolutamente continuo, la función de densidad de probabilidad no representa una probabilidad y, por tanto, puede valer más de 1. No obstante, en el caso absolutamente continuo, $f(x)dx$ se interpreta como la probabilidad infinitesimal de que la variable aleatoria X alcance valores dentro del intervalo $(x, x + dx)$ (véase la figura 1.1).

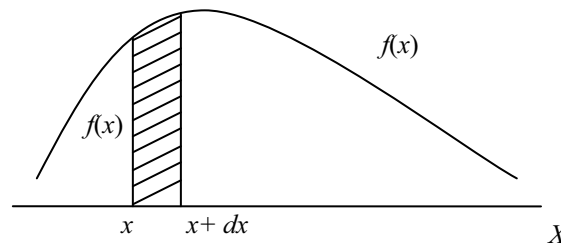


Figura 1.1

En el discreto, la probabilidad de que X tome un valor cualquiera a es:

$$P(X = a) \geq 0$$

En el absolutamente continuo, la probabilidad de que X tome un valor cualquiera a es nula; es decir:

$$P(X = a) = P(a < X < a) = \int_a^a f(x) dx = 0$$

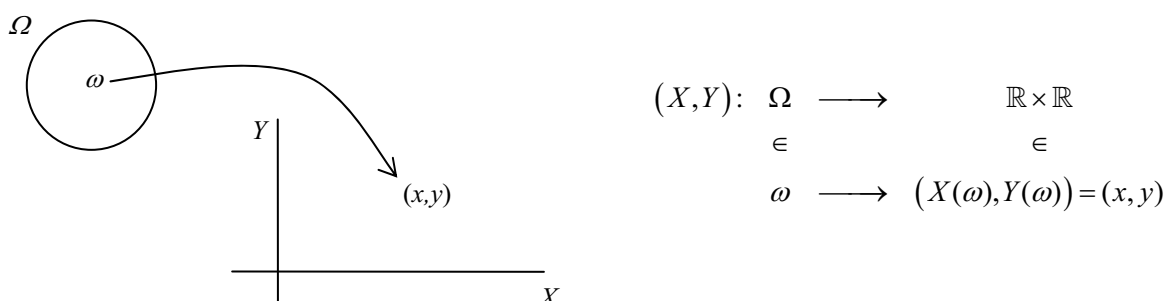
En el caso discreto, nos valemos de puntos para introducir la probabilidad, mientras que en el absolutamente continuo, nos valemos de intervalos, debido a ser nula la probabilidad en cada punto.

En el discreto, cualquier probabilidad es la suma de probabilidades asociadas a puntos. En el absolutamente continuo, cualquier probabilidad es una integral definida, asociada a un intervalo, verificándose:

$$P(a < X \leq b) = P(a < X < b) = P(a \leq X < b) = P(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(x) dx$$

1.5 Distribución conjunta de dos variables aleatorias

Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidades y, X e Y dos variables aleatorias definidas sobre él. Cada una de ellas es una aplicación de Ω en $\mathbb{R}$; por tanto, el par aleatorio (X, Y) permite definir una aplicación de Ω en el producto cartesiano $\mathbb{R} \times \mathbb{R}$ como sigue:



Se define una función de dos variables, que es la función de distribución conjunta del par aleatorio (X, Y) , de la siguiente manera:

$$F(x, y) = P\{\omega \in \Omega \mid [X(\omega) \leq x] \cap [Y(\omega) \leq y]\}$$

o en forma resumida:

$$F(x, y) = P\{[X \leq x] \cap [Y \leq y]\}$$

La función de distribución $F(x, y)$ tiene las siguientes propiedades:

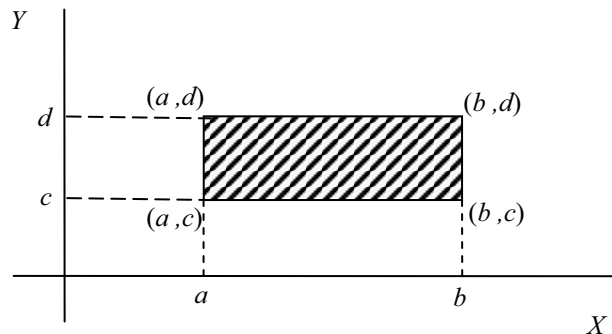
1) $0 \leq F(x, y) \leq 1$

2) $\lim_{\substack{x \rightarrow -\infty \\ o \\ y \rightarrow -\infty}} F(x, y) = 0 \quad \text{y} \quad \lim_{\substack{x \rightarrow +\infty \\ e \\ y \rightarrow +\infty}} F(x, y) = 1$

3) $\lim_{\substack{x \text{ fija} \\ y \rightarrow +\infty}} F(x, y) = F_1(x)$ siendo $F_1(x)$ la función de distribución marginal de la variable aleatoria X .

$\lim_{\substack{y \text{ fija} \\ x \rightarrow +\infty}} F(x, y) = F_2(y)$ siendo $F_2(y)$ la función de distribución marginal de la variable aleatoria Y .

4) $P\{[a < X \leq b] \cap [c < Y \leq d]\} = F(b, d) - F(a, d) - F(b, c) + F(a, c)$



Existen dos casos diferenciados: el caso discreto y el caso absolutamente continuo.

1.5.1 Distribuciones bivariantes discretas

Supóngase que X e Y son ambas discretas, o sea:

$$X: x_1, x_2, \dots, x_n, \dots$$

$$Y: y_1, y_2, \dots, y_m, \dots$$

Entonces el par aleatorio (X, Y) tomará los valores (x_i, y_j) .

La función de densidad conjunta en el caso discreto, se define así:

$$f(x_i, y_j) = P([X = x_i] \cap [Y = y_j]) = p_{ij} = P(x_i, y_j)$$

La función de densidad conjunta verifica las siguientes propiedades:

1) $0 \leq f(x_i, y_j) \leq 1$

2) $\sum_i \sum_j p_{ij} = 1$

3) $F(x, y) = \sum_{\substack{x_i \leq x \\ y_j \leq y}} f(x_i, y_j)$ siendo $F(x, y)$ la función de distribución conjunta del par aleatorio (X, Y) .

1.5.2 Funciones de densidad marginales en el caso discreto

Conociendo la distribución conjunta del par (X, Y) es posible determinar las distribuciones de X e Y por separado, denominadas *distribuciones marginales*.

La función de densidad marginal de la variable aleatoria X viene dada por:

$$f_1(x_i) = P[X = x_i] = \sum_j p_{ij}$$

En efecto,

$$[X = x_i] = \bigcup_j ([X = x_i] \cap [Y = y_j])$$

Los sucesos de la forma $[X = x_i] \cap [Y = y_j]$ son mutuamente excluyentes, por tanto:

$$P[X = x_i] = \sum_j P([X = x_i] \cap [Y = y_j]) = \sum_j p_{ij}$$

Análogamente, la función de densidad marginal de la variable aleatoria Y viene dada por:

$$f_2(y_j) = P[Y = y_j] = \sum_i p_{ij}$$

Ejemplo 1.3

Supónganse las variables aleatorias discretas X e Y con función de probabilidad conjunta dada por la siguiente tabla:

$X \backslash Y$	$y_1=4$	$y_2=5$
$x_1=3$	0,17	0,10
$x_2=10$	0,13	0,30
$x_3=12$	0,25	0,05

Se pide:

- Hallar las funciones de densidad marginales.
- Hallar los valores de la función de distribución conjunta $F(x_2, y_1)$, $F(x_2, y_2)$ y $F(x_3, y_2)$.
- Determinar $f(10, 5)$ y $F(10, 5)$.
- Función de distribución marginal de $X: F_1(x_2)$.
- Función de distribución marginal de $Y: F_2(y_2)$.

Solución:

- a) La función de densidad marginal de X se obtendrá sumando, para cada valor de X , los valores de la función de densidad conjunta, resultando:

$$f_1(x_1) = f_1(3) = P[X = 3] = f(3, 4) + f(3, 5) = 0,17 + 0,10 = 0,27$$

$$f_1(x_2) = f_1(10) = P[X = 10] = f(10, 4) + f(10, 5) = 0,13 + 0,30 = 0,43$$

$$f_1(x_3) = f_1(12) = P[X = 12] = f(12, 4) + f(12, 5) = 0,25 + 0,05 = 0,30$$

Se verifica que:

$$f_1(x_1) + f_1(x_2) + f_1(x_3) = 0,27 + 0,43 + 0,30 = 1$$

La función de densidad marginal de Y se obtendrá sumando, para cada valor de Y , los valores de la función de densidad conjunta, resultando:

$$f_2(y_1) = f_2(4) = P[Y = 4] = f(3, 4) + f(10, 4) + f(12, 4) = 0,17 + 0,13 + 0,25 = 0,55$$

$$f_2(y_2) = f_2(5) = P[Y = 5] = f(3, 5) + f(10, 5) + f(12, 5) = 0,10 + 0,30 + 0,05 = 0,45$$

Se verifica que:

$$f_2(y_1) + f_2(y_2) = 0,55 + 0,45 = 1$$

b)

$$F(x_2, y_1) = F(10, 4) = P([X \leq 10] \cap [Y \leq 4]) = f(3, 4) + f(10, 4) = 0,17 + 0,13 = 0,30$$

$$\begin{aligned} F(x_2, y_2) &= F(10, 5) = P([X \leq 10] \cap [Y \leq 5]) = \\ &= f(3, 4) + f(3, 5) + f(10, 4) + f(10, 5) = 0,17 + 0,10 + 0,13 + 0,30 = 0,70 \end{aligned}$$

$$\begin{aligned} F(x_3, y_2) &= F(12, 5) = P([X \leq 12] \cap [Y \leq 5]) = \\ &= f(3, 4) + f(3, 5) + f(10, 4) + f(10, 5) + f(12, 4) + f(12, 5) = 1 \end{aligned}$$

c)

$$f(10'5, 5) = P([X = 10'5] \cap [Y = 5]) = P(\emptyset \cap [Y = 5]) = P(\emptyset) = 0$$

$$F(10'5, 5) = P([X \leq 10'5] \cap [Y \leq 5]) = f(3, 4) + f(3, 5) + f(10, 4) + f(10, 5) = 0,7$$

d) Función de distribución marginal $X : F_1(x_2)$

$$F_1(x_2) = F_1(10) = P[X \leq 10] = f_1(3) + f_1(10) = 0,17 + 0,10 + 0,13 + 0,30 = 0,7$$

e) Función de distribución marginal $Y : F_2(y_2)$

$$F_2(y_2) = F_2(5) = P[Y \leq 5] = f_2(4) + f_2(5) = 0,17 + 0,13 + 0,25 + 0,10 + 0,30 + 0,05 = 1$$

1.5.3 Distribuciones bivariantes absolutamente continuas

Se dice que el par aleatorio (X, Y) tiene una distribución conjunta bivalente absolutamente continua si existe una función $f(x, y)$, llamada función de densidad bivalente, tal que:

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv$$

Se verifica:

$$1) f(x, y) \geq 0$$

$$2) \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$$

$$3) \frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y) \text{ si } f \text{ es continua en } (x, y).$$

$$4) P([a < X < b] \cap [c < Y < d]) = \int_a^b \int_c^d f(x, y) dx dy$$

5) $f(x, y)$ no es exactamente una probabilidad, pero $f(x, y) dx dy$ se interpreta como la probabilidad infinitesimal sobre el rectángulo de vértices (x, y) , $(x+dx, y)$, $(x, y+dy)$ y $(x+dx, y+dy)$. Es decir:

$$f(x, y) dx dy = P([x < X < x + dx] \cap [y < Y < y + dy])$$

1.5.4. Funciones de densidad marginales en el caso absolutamente continuo

En el caso absolutamente continuo las funciones de densidad marginales son:

$$f_1(x) = \int_{-\infty}^{+\infty} f(x, y) dy$$

$$f_2(y) = \int_{-\infty}^{+\infty} f(x, y) dx$$

1.6 Independencia estocástica de variables aleatorias

Dos variables aleatorias, X e Y , son estocásticamente independiente si, y sólo si, la función de distribución conjunta es el producto de las funciones de distribución marginales. Esto se cumple en general, o sea, tanto en el caso discreto como en el caso absolutamente continuo.

$$X, Y \text{ variables aleatorias independientes} \Leftrightarrow F(x, y) = F_1(x) \cdot F_2(y)$$

En efecto:

$$F(x, y) = P([X \leq x] \cap [Y \leq y]) = P([X \leq x]) \cdot P([Y \leq y]) = F_1(x) \cdot F_2(y)$$

1.6.1 Independencia de dos variables aleatorias en el caso discreto

Dos variables aleatorias, X e Y , son estocásticamente independientes si, y sólo si, la función de densidad conjunta es el producto de las funciones de densidad marginales.

$$f(x_i, y_j) = f_1(x_i) \cdot f_2(y_j)$$

para todo par de números reales (x_i, y_j) .

En efecto,

$$f(x_i, y_j) = P([X = x_i] \cap [Y = y_j]) = P([X = x_i]) \cdot P([Y = y_j]) = f_1(x_i) \cdot f_2(y_j)$$

1.6.2 Independencia de dos variables aleatorias en el caso absolutamente continuo

Dos variables aleatorias, X e Y , son estocásticamente independientes si, y sólo si, la función de densidad conjunta es el producto de las funciones de densidad marginales.

$$f(x, y) = f_1(x) \cdot f_2(y)$$

En efecto, por la propiedad (3) de la sección 1.5.3.

$$\begin{aligned} f(x, y) &= \frac{\partial^2 F(x, y)}{\partial x \partial y} = \frac{\partial^2 [F_1(x) \cdot F_2(y)]}{\partial x \partial y} \\ &= \frac{\partial}{\partial y} \left\{ \frac{\partial}{\partial x} [F_1(x) \cdot F_2(y)] \right\} = \frac{\partial}{\partial y} [F_2(y)] \cdot \frac{\partial}{\partial x} [F_1(x)] = \\ &= f_2(y) \cdot f_1(x) = f_1(x) \cdot f_2(y) \end{aligned}$$

1.7 Función de distribución condicional de X en el caso discreto

La función de distribución condicional de X en el caso discreto se define así:

$$G(x_i | y_j) = \frac{F(x_i, y_j)}{F_2(y_j)}$$

con $F_2(y_j) > 0$.

Es aquella que asigna a cada número real x_i la probabilidad de que X sea menor o igual que x_i , bajo la condición de que Y sea menor o igual que y_j , es decir,

$$G(x_i | y_j) = P([X \leq x_i] | [Y \leq y_j]) = \frac{P([X \leq x_i] \cap [Y \leq y_j])}{P([Y \leq y_j])} = \frac{F(x_i, y_j)}{F_2(y_j)}$$

1.8 Función de distribución condicional de Y en el caso discreto

Viene dada por la expresión siguiente:

$$H(y_j | x_i) = \frac{F(x_i, y_j)}{F_1(x_i)}$$

con $F_1(x_i) > 0$.

En efecto:

$$\begin{aligned} H(y_j | x_i) &= P([Y \leq y_j] | [X \leq x_i]) = \frac{P([Y \leq y_j] \cap [X \leq x_i])}{P([X \leq x_i])} = \\ &= \frac{P([X \leq x_i] \cap [Y \leq y_j])}{P([X \leq x_i])} = \frac{F(x_i, y_j)}{F_1(x_i)} \end{aligned}$$

1.9 Función de densidad condicional de X en el caso discreto

Viene dada por la siguiente expresión:

$$g(x_i | y_j) = g(x_i | Y = y_j) = \frac{f(x_i, y_j)}{f_2(y_j)}$$

con $f_2(y_j) > 0$.

En efecto,

$$g(x_i | y_j) = P([X = x_i] | [Y = y_j]) = \frac{P([X = x_i] \cap [Y = y_j])}{P([Y = y_j])} = \frac{f(x_i, y_j)}{f_2(y_j)}$$

1.10 Función de densidad condicional de Y en el caso discreto

Viene dada por la siguiente expresión:

$$h(y_j | x_i) = h(y_j | X = x_i) = \frac{f(x_i, y_j)}{f_1(x_i)}$$

con $f_1(x_i) > 0$.

En efecto,

$$h(y_j | x_i) = P([Y = y_j] | [X = x_i]) = \frac{P([Y = y_j] \cap [X = x_i])}{P([X = x_i])} = \frac{P([X = x_i] \cap [Y = y_j])}{P([X = x_i])} = \frac{f(x_i, y_j)}{f_1(x_i)}$$

1.11 Función de densidad condicional de X en el caso absolutamente continuo

Es una función de x , que viene dada por la siguiente expresión:

$$g(x | y_0) = \frac{f(x, y_0)}{f_2(y_0)}$$

con $f_2(y_0) > 0$.

1.12 Función de densidad condicional de Y en el caso absolutamente continuo

Es una función de y que viene dada por la siguiente expresión:

$$h(y | x_0) = \frac{f(x_0, y)}{f_1(x_0)}$$

con $f_1(x_0) > 0$.

Observación: La fórmula general de la función de densidad conjunta en el caso discreto es:

$$a) f(x_i, y_j) = f_1(x_i) \cdot h(y_j | x_i) = g(x_i | y_j) \cdot f_2(y_j)$$

En el caso particular de que las variables aleatorias X e Y son estocásticamente independientes, se sabe que se cumple:

$$b) f(x_i, y_j) = f_1(x_i) \cdot f_2(y_j)$$

Comparando a) y b) resulta: Si las variables aleatorias X e Y son estocásticamente independientes se verifica que

$$g(x_i | y_j) = f_1(x_i)$$

cualquiera que sea y_j y

$$h(y_j | x_i) = f_2(y_j)$$

cualquiera que sea x_i .

O sea, si las variables aleatorias X e Y son estocásticamente independientes se verifica que cada distribución condicional de X , para cada uno de los valores posibles de Y , es igual que la distribución marginal de X y que cada distribución condicional de Y , para cada uno de los valores posibles de X , es igual que la distribución marginal de Y .

1.13 Esperanza matemática

1.13.1 Sentido probabilístico de la esperanza matemática

Supóngase que se han realizado N pruebas, en las cuales la variable aleatoria X ha tomado f_1 veces el valor x_1 , f_2 veces el valor x_2 , ..., f_n el valor x_n ; además, $f_1 + f_2 + \dots + f_n = N$. En este caso, la suma de todos los valores tomados por X , es igual a

$$x_1 f_1 + x_2 f_2 + \dots + x_n f_n.$$

Se halla la media aritmética $\bar{X}$ de todos los valores tomados por la variable aleatoria, para lo cual se divide la suma obtenida por el número total de pruebas:

$$\bar{X} = \frac{x_1 f_1 + x_2 f_2 + \dots + x_n f_n}{N} \quad (1.1)$$

La relación f_1/N es la frecuencia relativa f_{R_1} del valor x_1 , f_2/N es la frecuencia relativa f_{R_2} de x_2 , etc., por tanto, la expresión (1.1) se puede escribir como:

$$\bar{X} = x_1 f_{R_1} + x_2 f_{R_2} + \dots + x_n f_{R_n} \quad (1.2)$$

Supóngase que el número de pruebas es bastante grande. Entonces la frecuencia relativa es aproximadamente igual a la probabilidad de aparición del suceso (esto se demostrará posteriormente).

Sustituyendo en (1.2) las frecuencias relativas por las correspondientes probabilidades, se obtiene:

$$\bar{X} \approx x_1 P[X = x_1] + x_2 P[X = x_2] + \dots + x_n P[X = x_n].$$

El segundo miembro de esta igualdad aproximada es la esperanza de la variable aleatoria X , o sea: $E(X)$.

El sentido probabilístico del resultado obtenido es: La esperanza matemática es aproximadamente igual (tanto más exacto, cuanto mayor es el número de pruebas) a la media aritmética de los valores observados de la variable aleatoria.

Dicho de otra forma: La esperanza matemática es una generalización del concepto de media aritmética.

1.13.2 Esperanza matemática de una variable aleatoria discreta

Supóngase que X es una variable aleatoria que alcanza los valores: $x_1, x_2, \dots, x_n, \dots$ con las probabilidades: $f(x_1), f(x_2), \dots, f(x_n), \dots$ siendo

$$f(x_i) = P[X = x_i]$$

Se definirá como *esperanza matemática* de la variable aleatoria X a la expresión:

$$E(X) = \sum_{i=1}^{n(\infty)} x_i P[X = x_i] = \sum_{i=1}^{n(\infty)} x_i f(x_i) \quad (1.3)$$

Si la variable aleatoria X toma infinitos valores, la expresión de $E(X)$ es una serie.

Si la serie es convergente, la $E(X)$ existe. Si la serie es divergente u oscilante, la esperanza no existe (no está definida).

A la esperanza de una variable aleatoria X también se la llama valor esperado o valor medio de la variable aleatoria X .

Ejemplo 1.4 (continuación de 1.1)

¿Cuál es la ganancia esperada del juego?

Solución:

Para hallar la ganancia esperada del juego deberemos hallar $E(X)$.

$$E(X) = 0,60 \frac{2}{15} + 1 \frac{1}{15} + 1,1 \frac{2}{10} + 1,5 \frac{2}{5} + 2 \frac{1}{5} = 1,37 \text{ €}.$$

1.13.3 Esperanza matemática de una variable aleatoria absolutamente continua

Supóngase que se tiene una variable aleatoria X con función de densidad $f(x)$. La esperanza X se define así:

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) \cdot dx \quad (1.4)$$

Ejemplo 1.5 (continuación de 1.2)

¿Cuál es la esperanza matemática de X ?

Solución:

Hallaremos la esperanza de X a partir de su definición. Esto es

$$E(x) = \int_{-\infty}^{\infty} x f_X(x) dx.$$

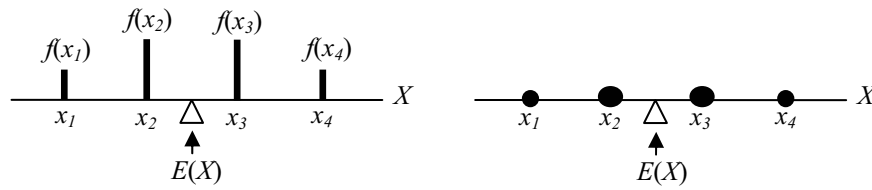
Resulta

$$E(x) = \frac{1}{12} \int_0^3 x(1+x^2) dx = \frac{1}{12} \left(\frac{1}{2} [x^2]_0^3 + \frac{1}{4} [x^4]_0^3 \right) = \frac{1}{12} \left(\frac{9}{2} + \frac{81}{4} \right) = \frac{33}{16}.$$

1.13.4 Significado de la esperanza

- 1) La esperanza es el valor medio teórico de los valores que alcanza la variable aleatoria. Es una *medida de centralización*.

- 2) La esperanza matemática de una variable aleatoria X se interpreta como el centro de gravedad de la distribución.



- 3) Si la variable aleatoria X describe la ganancia ó pérdida en un juego, la esperanza matemática indica la *ganancia media* que se espera obtener por cada jugada. Si la esperanza es cero, se dice que el juego es *equitativo*, si la esperanza es positiva se dice que el juego es *favorable* y si la esperanza es negativa se dice que el juego es *desfavorable*.

1.13.5 Propiedades de la esperanza matemática

Antes de establecer las propiedades de la esperanza enunciaremos los dos siguientes teoremas:

Teorema 1.1: Si X es una variable aleatoria con función de densidad $f(x)$ e $Y = g(X)$ es otra variable aleatoria construida a partir de la variable aleatoria X , mediante la función g , la $E(Y)$ en el caso discreto viene dada por:

$$E(Y) = \sum_i g(x_i) f(x_i)$$

y en el caso absolutamente continuo por:

$$E(Y) = \int_{-\infty}^{+\infty} g(x) f(x) dx$$

Teorema 1.2: Sean X e Y dos variables aleatorias con función de densidad conjunta $f(x, y)$. Construyamos la variable aleatoria $Z = g(X, Y)$.

En el caso discreto, la $E(Z)$ viene dada por:

$$E(Z) = \sum_i \sum_j g(x_i, y_j) \cdot f(x_i, y_j)$$

y en el caso absolutamente continuo por:

$$E(Z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) \cdot f(x, y) dx dy$$

Propiedades:

- 1) La esperanza matemática de una constante c es igual a la misma constante.
- 2) La esperanza matemática de la suma de varias variables aleatorias es igual a la suma de las esperanzas matemáticas de los sumandos. Para el caso de dos variables aleatorias X e Y , se tiene:

$$E(X + Y) = E(X) + E(Y).$$

Esta propiedad se cumple tanto si las variables aleatorias X e Y son estocásticamente independientes como si son dependientes.

3) Un factor constante se puede sacar fuera del signo de esperanza matemática:

$$E(cX) = c \cdot E(X)$$

4) Teniendo en cuenta las propiedades anteriores:

$$E(aX + bY + c) = aE(X) + bE(Y) + c$$

5) La esperanza matemática del producto de dos *variables aleatorias independientes* es igual al producto de sus esperanzas matemáticas:

$$E(X \cdot Y) = E(X) \cdot E(Y)$$

1.13.6 Esperanza matemática condicional

La distribución condicional de las probabilidades es una característica importante de la esperanza matemática condicional.

Se denomina *esperanza matemática condicional* de una variable aleatoria discreta Y para $X = x$ (x es un valor posible determinado de X) el producto de los valores posibles de Y por sus probabilidades condicionales:

$$E(Y|X = x) = \sum_j y_j h(y_j|x) \quad (1.5)$$

En el caso absolutamente continuo:

$$E(Y|X = x) = \int_{-\infty}^{+\infty} y h(y|x) dy \quad (1.6)$$

donde $h(y|x)$ es la función de densidad condicionada de la variable aleatoria Y cuando $X = x$.

Análogamente se determina la esperanza matemática condicional de la variable aleatoria X para $Y = y$ (siendo y un valor posible determinado Y).

Ejemplo 1.6

La distribución conjunta de dos variables aleatorias X e Y está dada por la siguiente tabla:

$X \backslash Y$	$y_1 = 3$	$y_2 = 6$
$x_1 = 1$	0,15	0,30
$x_2 = 3$	0,06	0,10
$x_3 = 4$	0,25	0,03
$x_3 = 8$	0,04	0,07

Determinar la esperanza matemática condicional de la variable aleatoria Y si $X = x_1 = 1$.

Solución:

Hallamos $f_1(x_1)$, para ello sumamos las probabilidades de la primera fila de la tabla:

$$f_1(x_1) = f_1(1) = 0,15 + 0,30 = 0,45$$

Hallamos la distribución condicional de las probabilidades de la variable aleatoria Y para $X = x_1 = 1$:

$$h(y_1|x_1) = h(3|1) = \frac{f(1,3)}{f_1(1)} = \frac{0,15}{0,45} = \frac{1}{3}$$

$$h(y_2|x_1) = h(6|1) = \frac{f(1,6)}{f_1(1)} = \frac{0,30}{0,45} = \frac{2}{3}$$

Hallamos la esperanza matemática condicional pedida aplicando la fórmula:

$$E(Y|X = x_1) = \sum_{j=1}^2 y_j h(y_j|x_1) = y_1 \cdot h(y_1|x_1) + y_2 \cdot h(y_2|x_1) = 3 \cdot h(3|1) + 6 \cdot h(6|1) = 3 \cdot \frac{1}{3} + 6 \cdot \frac{2}{3} = 5$$

1.14 Varianza de una variable aleatoria

La varianza mide la dispersión media de los valores de una variable respecto a su valor medio. En el caso de una muestra $x_1, x_2, \dots, x_n$, la dispersión de un valor x_i respecto a $\bar{X}$, se puede medir por $(x_i - \bar{X})^2$ y la media aritmética de esta dispersión es:

$$s^2 = \sum_{i=1}^n (x_i - \bar{X})^2 \cdot \frac{f_i}{N} = \sum_{i=1}^n (x_i - \bar{X})^2 \cdot f_R[X = x_i]$$

llamada varianza de la muestra. Su generalización para una variable aleatoria es:

Si X es una variable aleatoria con esperanza finita $E(X)$, se llama varianza de X , a la esperanza de la nueva variable $Y = [X - E(X)]^2$, es decir:

$$\sigma^2 = \text{var}(X) = E[X - E(X)]^2 \quad (1.5)$$

La raíz cuadrada positiva de $\text{var}(X)$ se llama desviación típica o desviación estándar y se indica por σ .

Observación: El número $\text{var}(X)$ está expresado en unidades cuadradas de X . Esto es, si X se mide en minutos, entonces $\text{var}(X)$ está expresada en $(\text{minutos})^2$. Esta es una razón para considerar la desviación estándar que se expresa en las mismas unidades que X .

1.14.1 Varianza de una variable aleatoria discreta

Supóngase que X es una variable aleatoria que alcanza los valores: $x_1, x_2, \dots, x_n, \dots$ con las probabilidades:

$$f(x_1), f(x_2), \dots, f(x_n), \dots$$

siendo

$$f(x_i) = P[X = x_i]$$

La varianza de la variable aleatoria X viene dada por la expresión:

$$\text{var}(X) = \sum_{i=1}^{n(\infty)} [x_i - E(X)]^2 \cdot P[X = x_i] = \sum_{i=1}^{n(\infty)} [x_i - E(X)]^2 \cdot f(x_i) \quad (1.6)$$

Ejemplo 1.7 (continuación de 1.1):

Hallar la varianza y la desviación estándar.

Solución:

De la definición

$$\text{var}(X) = (0,60 - 1,37)^2 \frac{2}{15} + (1 - 1,37)^2 \frac{1}{15} + (1,1 - 1,37)^2 \frac{2}{10} + (1,5 - 1,37)^2 \frac{2}{5} + (2 - 1,37)^2 \frac{1}{5} = 0,18$$

Finalmente, la desviación estándar es

$$\sigma(X) = \sqrt{\text{var}(X)} = \sqrt{0,18} = 0,43$$

1.14.2 Varianza de una variable aleatoria absolutamente continua

Supóngase que se tiene una variable aleatoria X absolutamente continua con función de densidad $f(x)$.

La varianza de la variable aleatoria X viene dada por la expresión:

$$\text{var}(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 \cdot f(x) \cdot dx \quad (1.7)$$

1.14.3 Propiedades de la varianza:

1) La varianza es siempre positiva o nula dado que $[X - E(X)]^2 \geq 0$.

2) La varianza se puede obtener también mediante la fórmula:

$$\text{var}(X) = E(X^2) - E^2(X) \quad (1.8)$$

3) La varianza de una constante es nula

$$\text{var}(c) = 0 \quad (1.9)$$

4) Si c es una constante:

$$\text{var}(X + c) = \text{var}(X) \quad (1.10)$$

5) Si c es una constante:

$$\text{var}(cX) = c^2 \text{var}(X) \quad (1.11)$$

6) Si X e Y son variables aleatorias estocásticamente independientes entonces:

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) \quad (1.12)$$

y

$$\text{var}(X - Y) = \text{var}(X) + \text{var}(Y) \quad (1.13)$$

7) Si X e Y son variables aleatorias cualesquiera entonces:

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2 \text{cov}(X, Y) \quad (1.14)$$

y, de forma equivalente,

$$\text{var}(X - Y) = \text{var}(X) + \text{var}(Y) - 2 \text{cov}(X, Y) \quad (1.15)$$

8) Si X e Y son dos variables aleatorias estocásticamente independientes entonces:

$$\text{var}(aX + bY + c) = a^2 \text{var}(X) + b^2 \text{var}(Y) \quad (1.16)$$

9) Si X e Y son dos variables aleatorias cualesquiera entonces:

$$\text{var}(aX + bY) = a^2 \text{var}(X) + b^2 \text{var}(Y) + 2ab \text{cov}(X, Y) \quad (1.17)$$

y, de forma equivalente.

$$\text{var}(aX - bY) = a^2 \text{var}(X) + b^2 \text{var}(Y) - 2ab \text{cov}(X, Y) \quad (1.18)$$

Ejemplo 1.8 (continuación del 1.2):

Hallar la varianza y la desviación estándar.

Solución:

Para hallar la varianza, tendremos en cuenta la igualdad

$$\text{var}(X) = E(X^2) - E^2(X)$$

Hallemos $E(X^2)$.

$$E(X^2) = \frac{1}{12} \int_0^3 x^2 (1 + x^2) dx = \frac{1}{12} \left(\frac{1}{3} [x^3]_0^3 + \frac{1}{5} [x^5]_0^3 \right) = \frac{1}{12} \left(\frac{27}{3} + \frac{243}{5} \right) = \frac{24}{5}$$

Por tanto,

$$\text{var}(X) = \frac{24}{5} - \left(\frac{33}{16} \right)^2 = \frac{699}{1280}$$

Finalmente, la desviación estándar es

$$\sigma(X) = \sqrt{\text{var}(X)} = \sqrt{0,546} = 0,739$$

1.15 Teorema de Tchebychev

Sea X una variable aleatoria con distribución cualquiera y con esperanza y varianza finitas. La probabilidad de que la variable aleatoria X caiga dentro de k desviaciones estándar de la media es al menos $\left(1 - \frac{1}{k^2}\right)$.

Esto es,

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

o en forma más resumida:

$$P\{|X - \mu| < k\sigma\} \geq 1 - \frac{1}{k^2}$$

Haciendo $h = k\sigma$, se puede escribir:

$$P\{|X - \mu| < h\} \geq 1 - \frac{\sigma^2}{h^2}$$

siendo μ la esperanza matemática de la variable aleatoria X y σ^2 su varianza.

El teorema de Tchebychev también se puede expresar así:

$$P\{|X - \mu| \geq h\} \leq \frac{\sigma^2}{h^2}$$

es decir, la probabilidad de que la variable aleatoria X se aparte de su esperanza en una cantidad mayor o igual que h ($h > 0$) es menor o igual que σ^2 / h^2 .

Únicamente se demostrará el caso absolutamente continuo, dejando el caso discreto como ejercicio para el lector.

Demostración:

A partir de la definición previa de la varianza de X , se puede escribir:

$$\begin{aligned} \sigma^2 &= E(X - \mu)^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx = \\ &= \int_{-\infty}^{\mu - k\sigma} (x - \mu)^2 f(x) dx + \int_{\mu - k\sigma}^{\mu + k\sigma} (x - \mu)^2 f(x) dx + \int_{\mu + k\sigma}^{+\infty} (x - \mu)^2 f(x) dx \geq \\ &\geq \int_{-\infty}^{\mu - k\sigma} (x - \mu)^2 f(x) dx + \int_{\mu + k\sigma}^{+\infty} (x - \mu)^2 f(x) dx \end{aligned}$$

ya que la segunda de las tres integrales no es negativa. Como $|x - \mu| \geq k\sigma$ siempre que $x \geq \mu + k\sigma$ ó $x \leq \mu - k\sigma$, se tiene $(x - \mu)^2 \geq k^2 \sigma^2$ en las dos integrales restantes. De aquí se desprende que:

$$\begin{aligned} \sigma^2 &\geq \int_{-\infty}^{\mu - k\sigma} k^2 \sigma^2 f(x) dx + \int_{\mu + k\sigma}^{+\infty} k^2 \sigma^2 f(x) dx \\ k^2 \sigma^2 &\left(\int_{-\infty}^{\mu - k\sigma} f(x) dx + \int_{\mu + k\sigma}^{+\infty} f(x) dx \right) \leq \sigma^2 \end{aligned}$$

Dividiendo los dos miembros de la desigualdad por $k^2\sigma^2$ se tiene:

$$\int_{-\infty}^{\mu-k\sigma} f(x)dx + \int_{\mu+k\sigma}^{+\infty} f(x)dx \leq \frac{1}{k^2}$$

Multiplicando por (-1) los dos miembros de la desigualdad:

$$-\int_{-\infty}^{\mu-k\sigma} f(x)dx - \int_{\mu+k\sigma}^{+\infty} f(x)dx \geq -\frac{1}{k^2}$$

Sumando 1 a los dos miembros de la desigualdad:

$$1 - \int_{-\infty}^{\mu-k\sigma} f(x)dx - \int_{\mu+k\sigma}^{+\infty} f(x)dx \geq 1 - \frac{1}{k^2}$$

Teniendo presente que

$$\int_{-\infty}^{\mu-k\sigma} f(x)dx + \int_{\mu-k\sigma}^{\mu+k\sigma} f(x)dx + \int_{\mu+k\sigma}^{+\infty} f(x)dx = 1$$

resulta:

$$\int_{\mu-k\sigma}^{\mu+k\sigma} f(x)dx = P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

que es lo que se quería comprobar.

Para $k=2$ el teorema establece que la variable aleatoria X tiene por lo menos una probabilidad de $1 - (1/2)^2 = 3/4$ de caer entre dos desviaciones estándar de la media. Es decir, tres cuartas partes o más de las observaciones de cualquier distribución se encuentran en el intervalo $\mu \pm 2\sigma$. Análogamente el teorema afirma que al menos $8/9$ de las observaciones de cualquier distribución se encuentran en el intervalo $\mu \pm 3\sigma$.

El teorema de Tchebychev proporciona la cota inferior de la probabilidad, pero no su valor exacto. Se sabe que la probabilidad de que una variable aleatoria caiga dentro de dos desviaciones estándar alrededor de su media no puede ser menor de $3/4$, pero no se sabe el valor exacto de la probabilidad. En el caso de que se conozca la distribución de probabilidad que sigue la variable aleatoria X se podrá determinar el valor exacto de la probabilidad.

Ejemplo 1.9

Una variable aleatoria X tiene una media $\mu=100$, una varianza $\sigma^2=16$, y una distribución de probabilidad desconocida. Acotar:

- $P(84 < X < 116)$.
- $P\{|X - 100| \geq 8\}$

Solución:

$$\text{a) } P(84 < X < 116) = P[100 - (4) \cdot (4) < X < 100 + (4) \cdot (4)] \geq 1 - \frac{1}{4^2} = \frac{15}{16}$$

b)

$$\begin{aligned} P\{|X - 100| \geq 8\} &= 1 - P\{|X - 100| < 8\} = \\ &= 1 - P(-8 < X - 100 < 8) = \\ &= 1 - P(92 < X < 108) = \\ &= 1 - P[100 - (2) \cdot (4) < X < 100 + (2) \cdot (4)] \\ &\leq \frac{1}{4} \end{aligned}$$

ya que

$$P[100 - (2) \cdot (4) < X < 100 + (2) \cdot (4)] \geq \frac{3}{4}.$$

Multiplicando por (-1) los dos miembros de la desigualdad:

$$-P[100 - (2) \cdot (4) < X < 100 + (2) \cdot (4)] \leq -\frac{3}{4}$$

y finalmente sumando 1 a los dos lados miembros resulta:

$$1 - P[100 - (2) \cdot (4) < X < 100 + (2) \cdot (4)] \leq \frac{1}{4}$$

1.16 Momentos de una variable aleatoria X respecto del origen

Dada una variable aleatoria X , se llama *momento de orden r (respecto del origen)*, a la esperanza de la variable X^r :

$$m_r = E(X^r)$$

si existe.

1.16.1 Caso discreto

Sea X una variable aleatoria discreta con función de densidad $f(x)$. El momento de orden r (respecto del origen) de la variable aleatoria X es:

$$m_r = E(X^r) = \sum x^r f(x)$$

Se observa que el momento de 1^{er} orden (respecto del origen) es la esperanza ó media:

$$m_1 = E(X) = \sum x \cdot f(x) = \mu$$

1.16.2 Caso absolutamente continuo

Sea X una variable aleatoria continua con función de densidad $f(x)$. El momento de orden r (respecto del origen) de la variable aleatoria X es:

$$m_r = E(X^r) = \int_{-\infty}^{+\infty} x^r f(x) dx$$

El momento de 1^{er} orden (respecto del origen) es la esperanza ó media:

$$m_1 = E(X) = \int_{-\infty}^{+\infty} x f(x) dx = \mu$$

1.17 Momentos centrales de una variable aleatoria X

Dada una variable aleatoria X , se denomina *momento central de orden r (respecto de la esperanza de X)*, a la esperanza de la variable $[X - E(X)]^r$, o sea:

$$\mu_r = E[X - E(X)]^r$$

1.17.1 Caso discreto

Sea X una variable aleatoria discreta con función de densidad $f(x)$. El momento central de orden r de la variable aleatoria X viene dado por:

$$\mu_r = E[X - E(X)]^r = \sum [x - E(X)]^r f(x)$$

Obsérvese que el momento central de 2^o orden es la varianza X , representada por $\text{var}(X)$ y también por σ^2 :

$$\mu_2 = E[X - E(X)]^2 = \sum [x - E(X)]^2 f(x) = \sum (x - \mu)^2 f(x) = \text{var}(X) = \sigma^2$$

1.17.2 Caso absolutamente continuo

Sea X una variable aleatoria continua con función de densidad $f(x)$. El momento central de orden r de la variable aleatoria X viene dado por:

$$\mu_r = E[X - E(X)]^r = \int_{-\infty}^{+\infty} [x - E(X)]^r f(x) dx$$

El momento central de 2^o orden es la varianza de X :

$$\mu_2 = E[X - E(X)]^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx = \text{var}(X) = \sigma^2$$

1.18 Momento de dos variables aleatorias respecto del origen

Sea X_1 y X_2 dos variables aleatorias. Se denomina momento de orden $r_1 + r_2$ (respecto del origen) de las variables aleatorias X_1 y X_2 a la expresión:

$$m_{r_1 r_2} = E(X_1^{r_1} \cdot X_2^{r_2})$$

si existe la esperanza.

1.18.1 Caso discreto

Sean X_1 y X_2 dos variables aleatorias discretas con función de densidad conjunta $f(x_1, x_2)$. El momento de orden $r_1 + r_2$ (respecto del origen) de las variables aleatorias X_1 y X_2 es:

$$m_{r_1 r_2} = E(X_1^{r_1} \cdot X_2^{r_2}) = \sum_1 \sum_2 x_1^{r_1} x_2^{r_2} f(x_1, x_2)$$

1.18.2 Caso absolutamente continuo

Sean X_1 y X_2 dos variables aleatorias continuas con función de densidad conjunta $f(x_1, x_2)$. El momento de orden $r_1 + r_2$ (respecto del origen) de las variables aleatorias X_1 y X_2 es:

$$m_{r_1 r_2} = E(X_1^{r_1} \cdot X_2^{r_2}) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_1^{r_1} x_2^{r_2} f(x_1, x_2) dx_1 dx_2$$

1.19 Momentos centrales de dos variables aleatorias

Sean X_1 y X_2 dos variables aleatorias. Se denomina momento central de orden $r_1 + r_2$ (respecto de las esperanzas de X_1 y X_2) de las variables aleatorias X_1 y X_2 a la expresión:

$$\mu_{r_1 r_2} = E\left\{[X_1 - E(X_1)]^{r_1} \cdot [X_2 - E(X_2)]^{r_2}\right\}$$

1.19.1 Caso discreto

Sean X_1 y X_2 dos variables aleatorias discretas con función de densidad conjunta $f(x_1, x_2)$. El momento central de orden $r_1 + r_2$ de las variables aleatorias X_1 y X_2 es:

$$\begin{aligned} \mu_{r_1 r_2} &= E\left\{[X_1 - E(X_1)]^{r_1} \cdot [X_2 - E(X_2)]^{r_2}\right\} = \\ &= \sum_1 \sum_2 [x_1 - E(X_1)]^{r_1} [x_2 - E(X_2)]^{r_2} f(x_1, x_2) \end{aligned}$$

1.19.2 Caso absolutamente continuo

Sean X_1 y X_2 dos variables aleatorias continuas con función de densidad conjunta $f(x_1, x_2)$. El momento central de orden $r_1 + r_2$ de las variables aleatorias X_1 y X_2 es:

$$\begin{aligned}\mu_{r_1 r_2} &= E\left\{\left[X_1 - E(X_1)\right]^{r_1} \cdot \left[X_2 - E(X_2)\right]^{r_2}\right\} = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left[x_1 - E(X_1)\right]^{r_1} \left[x_2 - E(X_2)\right]^{r_2} f(x_1, x_2) dx_1 dx_2\end{aligned}$$

1.20 Covarianza de dos variables aleatorias X e Y

La covarianza de dos variables aleatorias X e Y es por definición el momento central de orden $1 + 1$. Es decir:

$$\text{cov}(X, Y) = \mu_{11} = E\left\{\left[X - E(X)\right] \cdot \left[Y - E(Y)\right]\right\} \quad (1.19)$$

La covarianza entre las variables aleatorias X e Y se representa por $\text{cov}(X, Y)$ y también por σ_{XY} .

1.20.1 Propiedades

La covarianza tiene las siguientes propiedades:

- 1) $\text{cov}(X, Y) = E(X \cdot Y) - E(X) \cdot E(Y)$
- 2) $\text{cov}(X, Y) = \text{cov}(Y, X)$
- 3) $\text{cov}(X + a, Y) = \text{cov}(X, Y)$
- 4) $\text{cov}(aX, Y) = a \text{cov}(X, Y)$
- 5) $\text{cov}(aX, bY) = ab \text{cov}(X, Y)$
- 6) $\text{cov}(X + Y, Z) = \text{cov}(X, Z) + \text{cov}(Y, Z)$
- 7) Combinando las propiedades anteriores:

$$\begin{aligned}\text{cov}(aX + bY + c, dZ + eW + f) &= \\ &= ad \text{cov}(X, Z) + ae \text{cov}(X, W) + bd \text{cov}(Y, Z) + be \text{cov}(Y, W)\end{aligned}$$

8) Si X e Y son variables aleatorias estocásticamente independientes su covarianza es cero. El recíproco no es cierto; es decir puede ser nula la covarianza de dos variables aleatorias y no ser éstas independientes.

La covarianza proporciona una medida del grado de dependencia entre X e Y . La covarianza tiene el inconveniente de que es proporcional a la magnitud de las variables aleatorias. Es decir, da el

grado de dependencia pero no estandarizado. Para dar este grado de dependencia de forma estándar se utiliza otro concepto que se llama *coeficiente de correlación de Pearson*, designado por ρ .

1.21 Coeficiente de correlación lineal de Pearson

Sean X e Y las variables aleatorias con momentos de segundo orden finitos y varianzas no nulas. Al valor:

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)}\sqrt{\text{var}(Y)}}$$

se le denomina coeficiente de correlación lineal de Pearson entre X e Y .

El coeficiente de correlación es una cantidad adimensional.

Teorema: Si X e Y son dos variables aleatorias estocásticamente independientes, entonces $\rho = 0$.

En efecto:

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)}\sqrt{\text{var}(Y)}}$$

Si X e Y son independientes, entonces por la propiedad (1) de la covarianza, $\text{cov}(X, Y) = 0$ y al ser cero la covarianza, el coeficiente de correlación ρ también será cero.

Observación: El recíproco del teorema no es verdadero. Esto es, se puede tener $\rho = 0$, y aún así X e Y no necesitan ser independientes. Si $\rho = 0$ se dice que las variables aleatorias X e Y están incorrelacionadas, pero podrían ser estocásticamente dependientes. Así los conceptos “no correlacionadas” e “independientes” no son, en general, equivalentes.

1.22 Desigualdad de Schwarz

Sean V y W dos variables aleatorias sobre un espacio de probabilidades $(\Omega, \mathcal{A}, P)$ con momentos de primer y segundo orden finitos. Se verifica:

$$[E(V \cdot W)]^2 \leq E(V^2) \cdot E(W^2)$$

y si

$$[E(V \cdot W)]^2 = E(V^2) \cdot E(W^2)$$

entonces $W = aV$ para una cierta constante a , es decir, entre las variables V y W existe una relación de dependencia lineal.

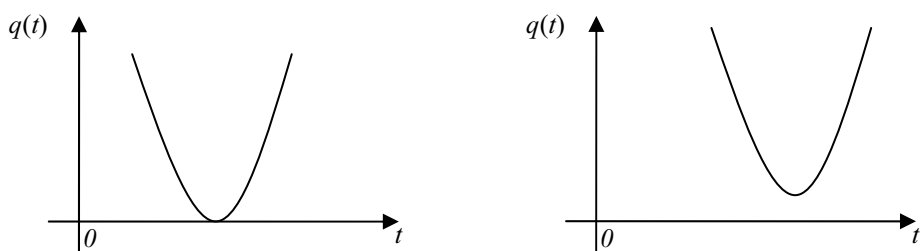
Demostración: Considérese la siguiente función de la variable real t :

$$q(t) = E[V + t \cdot W]^2.$$

Puesto que $[V + tW]^2 \geq 0$, se tiene que $q(t) \geq 0$ para todo t . Desarrollando se obtiene:

$$q(t) = E[V^2 + 2tVW + t^2W^2] = E(V^2) + 2tE(V \cdot W) + t^2E(W^2)$$

Así $q(t)$ es un polinomio de segundo grado en t . En general, si un polinomio de 2º grado $q(t) = at^2 + bt + c$ tiene la propiedad de que $q(t) \geq 0$ para todo t , significa que su gráfica corta el eje t en un sólo punto o en ninguno, como se indica en la figura siguiente.



Esto, a su vez, indica que el discriminante $b^2 - 4ac$ debe ser ≤ 0 , puesto que $b^2 - 4ac > 0$ significaría que tiene dos raíces reales distintas. Aplicando esta conclusión a la función $q(t)$ que se consideró anteriormente se obtiene:

$$\Delta = 4[E(V \cdot W)]^2 - 4E(V^2)E(W^2) \leq 0$$

Dividiendo por 4 resulta:

$$\begin{aligned} [E(V \cdot W)]^2 - E(V^2) \cdot E(W^2) &\leq 0 \\ [E(V \cdot W)]^2 &\leq E(V^2)E(W^2) \end{aligned}$$

que es la desigualdad de Schwarz.

Si el discriminante es cero, entonces para un solo valor $t = t_0$ se verifica $q(t) = 0$, es decir,

$$E[V + t_0 W]^2 = 0$$

pero esto implica que

$$V + t_0 W = 0,$$

o sea

$$W = -\frac{1}{t_0}V = aV$$

con $a = \frac{-1}{t_0}$.

1.23 Propiedades del coeficiente de correlación de Pearson

- 1) $-1 \leq \rho \leq 1$.
- 2) Si $\rho^2 = 1$, o bien, $\rho = \pm 1$, entre las variables aleatorias X e Y existe una relación lineal $Y = aX + b$.
- 3) $\rho(X, Y) = \rho(Y, X)$

4) $\rho(X, X) = 1$

5) $\rho(X, -X) = -1$

6) $\rho(aX + b, cY + d) = \rho(X, Y)$ si $a, c \neq 0$