

Resumen de Estadística

Maximiliano Illbele

8 de diciembre de 2011

Índice

1. Estadística descriptiva	5
1.1. Conceptos básicos y terminología	5
1.2. Población y muestra	5
1.3. Tipos de datos	6
1.4. Tabla de distribución de frecuencias	7
1.4.1. Histograma	7
1.5. Medidas de posición o tendencia central	11
1.6. Medidas de dispersión o variabilidad	13
1.7. Guía para la construcción de un gráfico de caja o box-plot . .	14
2. Probabilidad	20
2.1. Propiedades de un modelo probabilístico	22
2.2. Probabilidad condicional	24
2.2.1. Propiedades de la probabilidad condicional	25
2.3. Teorema de Bayes	26
2.4. Eventos independientes	27
2.5. Variables aleatorias	28
2.5.1. Clasificación de las variables aleatorias	28
3. Variables aleatorias discretas	29
3.1. Propiedades de una variable aleatoria discreta	30
3.2. Valor esperado de una variable aleatoria discreta	33
3.3. Varianza de una variable aleatoria discreta	33
3.4. Propiedades de la varianza y la esperanza de una V.A. discreta	33
3.5. Distribución Binomial	38
3.5.1. Función de distribución de probabilidad de una binomial	39
3.5.2. Esperanza de una binomial	40
3.6. Distribución de Poisson	42
3.6.1. Esperanza y varianza de una Poisson	43
3.6.2. Distribución de Poisson vista como una binomial . . .	45
3.7. Distribución hiper geométrica	46
3.7.1. Propiedades de la hiper geométrica	46
3.8. Distribución binomial negativa	48
4. Variables aleatorias continuas	49
4.1. Propiedades	49
4.2. Función de distribución acumulada	51
4.3. Distribución uniforme	54
4.4. Esperanza y varianza de una variable aleatoria continua . . .	55

4.4.1.	Esperanza y varianza de una uniforme	56
4.5.	Distribución normal	58
4.5.1.	Propiedades de la función de densidad	58
4.5.2.	Estandarización	62
4.6.	Distribución Gamma	64
4.7.	Distribución exponencial	68
4.7.1.	Distribución exponencial vista como una Gamma . . .	68
4.7.2.	Propiedades de la distribución exponencial	68
4.8.	Distribución χ^2	70
4.8.1.	Propiedades de la distribución χ^2	70
5.	Distribución de probabilidad conjunta	72
5.1.	Funciones marginales	72
5.2.	Función de densidad conjunta	73
5.3.	Interdependencia entre variables aleatorias	76
5.4.	Covarianza	79
5.4.1.	Propiedades de la covarianza	79
5.4.2.	Probabilidad conjunta y covarianza	80
5.5.	Definición de correlación	81
5.5.1.	Generalización	83
5.6.	Teorema Central del Límite T.C.L	88
5.7.	Aproximación normal a una binomial	88
6.	Estimación puntual	93
6.1.	Métodos para la estimación de los parámetros de la distribución	97
6.1.1.	Método de los Momentos	97
6.1.2.	Método de Máxima Verosimilitud	99
6.1.3.	Propiedades de los E.M.V	103
7.	Intervalos de confianza basados en una sola muestra	104
7.1.	Método general para obtener un IC para θ	104
7.2.	IC para $\theta = \mu$ con muestras suficientemente grandes	106
7.3.	Distribución t de Student con n grados de libertad	109
7.3.1.	Características de la t de Student	109
7.4.	Intervalo de confianza para la varianza	111
8.	Prueba de hipótesis basada en una muestra	112
8.1.	Componentes de una prueba de hipótesis	112
8.1.1.	Tipos de de hipótesis alternativas (H_a)	112
8.1.2.	Región de rechazo de la prueba (RR)	112
8.1.3.	Tipos de errores de una prueba de hipótesis	112

8.1.4. Nivel de significancia	113
8.2. Pasos a seguir para hacer una prueba de hipótesis	113
8.3. Prueba de hipótesis para la media poblacional	113
8.4. Prueba de hipótesis para la varianza poblacional	118
8.5. Prueba de hipótesis para la proporción poblacional	118
8.6. P-Valor	120
9. Inferencia basada en dos muestras	121

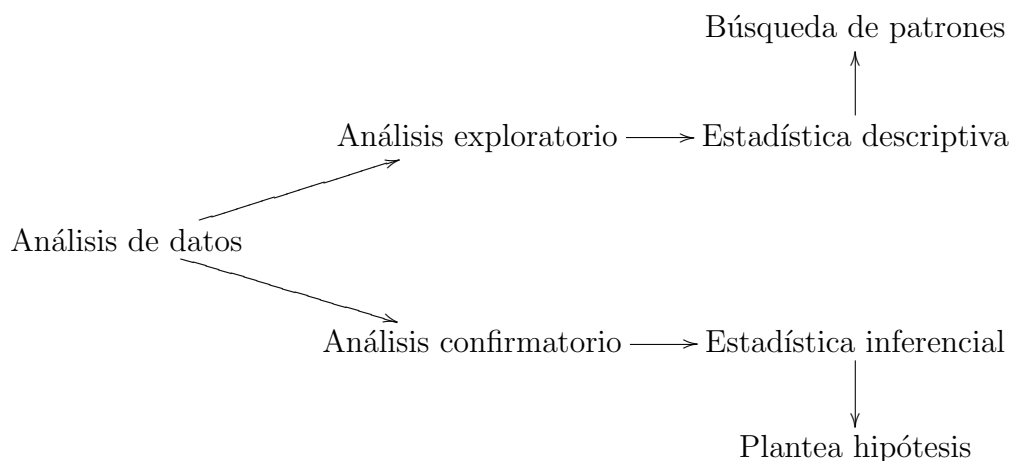
1. Estadística descriptiva

1.1. Conceptos básicos y terminología

La disciplina de la estadística enseña cómo razonar de manera lógica y a tomar decisiones informadas en presencia de incertidumbre y variación.

Podemos distinguir dos ramas principales de la estadística según como trabajen con los datos generados por cualquier tipo de investigación.

- Estadística descriptiva: se encarga de organizar los datos y de resumir la información.
- Estadística inferencial: se encarga de extraer conclusiones según las hipótesis planteadas.



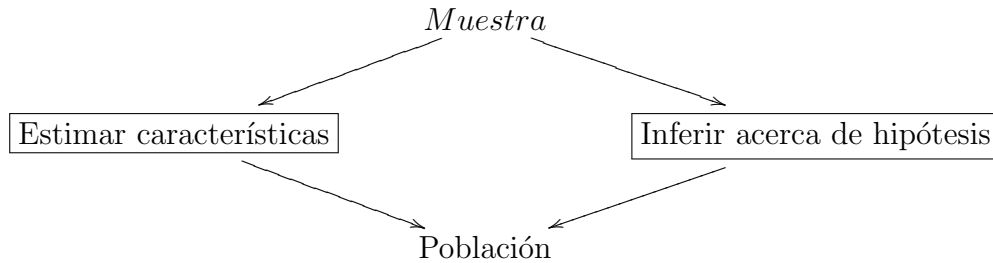
1.2. Población y muestra

Población:

- Colección de elementos o sujetos de interés.
- Puede ser finita o infinita.

Muestra:

- Subconjunto elegido al azar de la población.
- Tamaño muestral n .



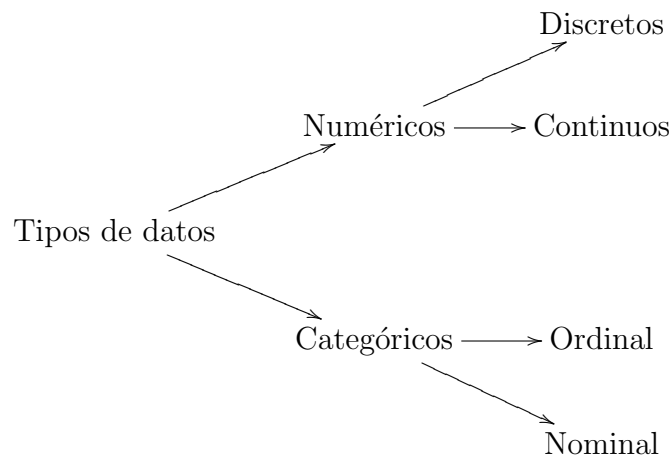
1.3. Tipos de datos

Numéricos

- Discretos: (determinados valores) como ser el número de hermanos o el número de accidentados en un choque de motos.
- Continuos: (valores en un intervalo) como ser la concentración de glucosa en sangre, o el nivel de $PH[0, 14]$.

Categóricos

- Ordinal: (orden, jerarquías) como ser el estado que atravieza una enfermedad (severo, moderado, suave) o el máximo nivel educativo alcanzado (primario, secundario, terciario, universitario).
- Nominal: (sin orden) como ser el grupo sanguíneo de una persona, la religión o el estado civil.



Dijimos que la estadística descriptiva es la encargada de organizar y resumir la información de los datos, esto lo hace, teniendo en cuenta el conjunto de datos, seleccionando el método más adecuado:

- Tablas de distribución de frecuencias.
- Medidas de posición o tendencia central.
- Medidas de dispersión o variabilidad.
- Gráficos.

1.4. Tabla de distribución de frecuencias

Método:

1. Tomar un intervalo que contenga al conjunto de datos.
2. Dividir el intervalo en k sub-intervalos de clase (IC), adyacentes y disjuntos.
3. Contar el número de observaciones en cada intervalo (FA)¹.
4. Calcular las (FR)² como el cociente entre las FA y el tamaño muestral (n) en cada uno de los k intervalos. (Generalmente $k = \sqrt{n}$).
5. Se puede agregar:
 - Marca de clase: punto medio de cada intervalo de clase.
 - Frecuencia absoluta acumulada. (FAA)
 - Frecuencia relativa acumulada. (FRA)

1.4.1. Histograma

- Gráfico de mayor difusión, es la representación gráfica de la tabla de distribución de frecuencias.
- ¿Cómo hacerlo?
 - En una recta horizontal marcar los k intervalos de clase (IC).
 - Sobre cada intervalo trazar un rectángulo cuya área sea proporcional al número de observaciones del mismo.

¹Frecuencias absolutas

²Frecuencias relativas

- ¿Cómo elegir la altura de los rectángulos?

$$Altura = \frac{FR}{Long(IC)}$$

- Los intervalos no tienen que tener la misma longitud, si los IC son de igual longitud entonces las alturas de los rectángulos son proporcionales a las frecuencias absolutas o a las frecuencias relativas. Luego comparar dos IC se reduce a ver sus alturas.

Si los intervalos son de distintas longitudes, para comparar 2 IC debemos comparar sus áreas y no sus alturas.

- **Observaciones:**
 - $\sum_{i=1}^k FA_i = n$
 - $\sum_{i=1}^k FR_i = 1$

Ejemplo

Para decidir el número de cajas necesarias para trabajar en un supermercado, se requiere tener información sobre el tiempo, medido en minutos, requerido para atender a los clientes.

Para tal fin, se tomó una muestra al azar a 60 clientes y se midió el tiempo que se demoró en atenderlos.

Los datos obtenidos, ordenados de menor a mayor, fueron:

0,20	0,20	0,30	0,30	0,30	0,40	0,40	0,40	0,50	0,50
0,60	0,60	0,60	0,60	0,70	0,70	0,70	0,80	0,80	0,80
0,80	0,90	0,90	1,00	1,00	1,10	1,10	1,10	1,10	1,10
1,10	1,10	1,20	1,20	1,20	1,30	1,30	1,30	1,40	1,40
1,60	1,60	1,70	1,70	1,80	1,80	1,80	1,80	1,90	1,90
2,10	2,20	2,30	2,50	2,80	3,10	3,10	3,60	4,50	5,20

Primero: Número de intervalos de clase: $k = \sqrt{60} = 7,75 \sim 8$.

Segundo: Elegir la longitud de los intervalos de clase.

Si queremos una partición disjunta del intervalo $[0,2; 5,2]$ en 8 *sub-intervalos* de igual longitud (L).

Ésta debe ser igual a: $L = \frac{5,2-0,2}{8} = 0,625$

Tabla de distribución de frecuencias

IC	FA	FR
[0.2, 0.825)	21	$21/60 = 0.35$
[0.825, 1.45)	19	$19/60 = 0.32$
[1.45, 2.075)	10	$10/60 = 0.17$
[2.075, 2.7)	4	$4/60 = 0.07$
[2.7, 3.325)	3	$3/60 = 0.05$
[3.325, 3.95)	1	$1/60 = 0.02$
[3.95, 4.575)	1	$1/60 = 0.02$
[4.575, 5.2]	1	$1/60 = 0.02$
n = 60		$\Sigma = 1$

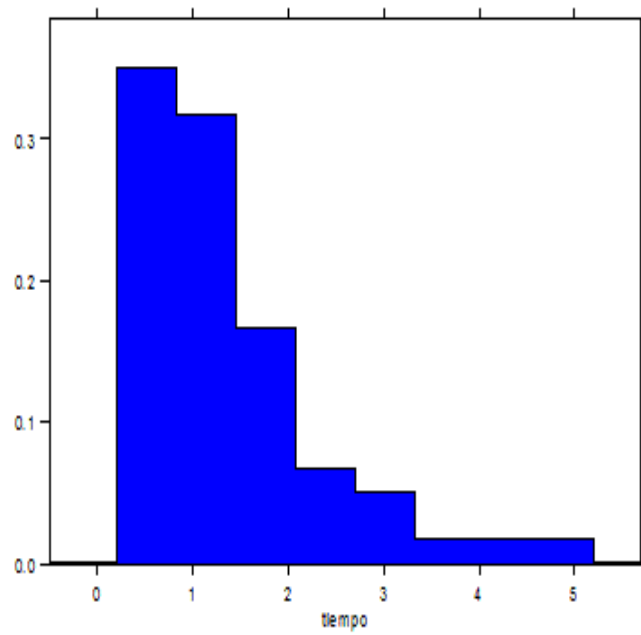


Figura 1: Histograma asociado

Ejemplo

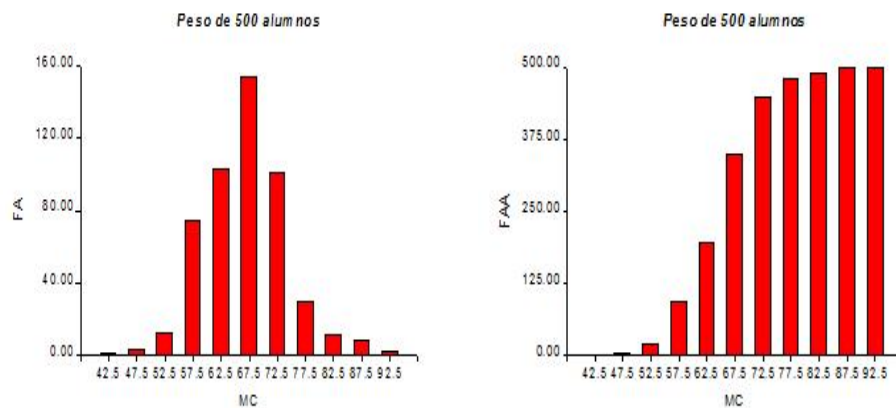
Distribución del peso (x) en Kg de una muestra de 500 alumnos varones de una universidad.

IC	FA	FAA	FR	FRA	%	Marca de clase
$40 < x \leq 45$	1	1	0.002	0.002	0.2	42.5
$45 < x \leq 50$	3	4	0.006	0.008	0.6	47.5
$50 < x \leq 55$	12	16	0.024	0.032	2.4	52.5
$55 < x \leq 60$	75	91	0.150	0.182	15.0	57.5
$60 < x \leq 65$	103	194	0.206	0.388	20.6	62.5
$65 < x \leq 70$	155	349	0.310	0.698	31.0	67.5
$70 < x \leq 75$	101	450	0.202	0.900	20.2	72.5
$75 < x \leq 80$	29	479	0.058	0.958	5.8	77.5
$80 < x \leq 85$	11	490	0.022	0.980	2.2	82.5
$85 < x \leq 90$	8	498	0.016	0.996	1.6	87.5
$90 < x \leq 95$	2	500	0.004	1.000	0.4	92.5
Total	500	500	1.000	1.000	100.0	-

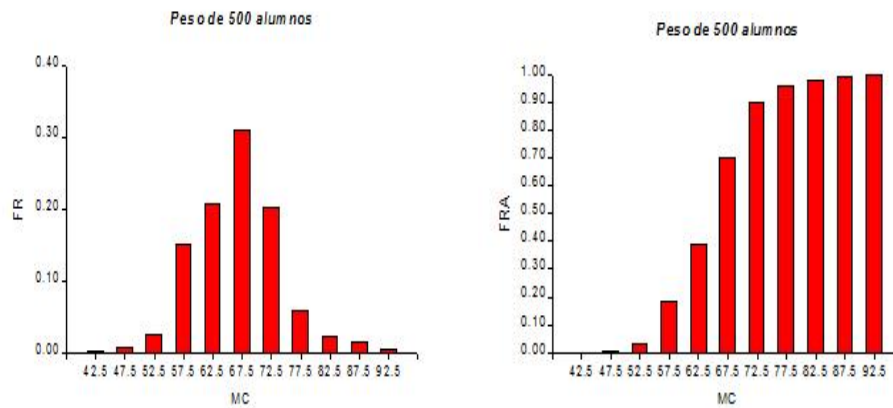
FAA= Frecuencias absolutas acumuladas

FRA= Frecuencias relativas acumuladas

Histogramas de las frecuencias absolutas



Histogramas de las frecuencias relativas



1.5. Medidas de posición o tendencia central

- Media muestral o promedio muestral

Sean x_1, \dots, x_n los datos obtenidos en la muestra

- La definimos como: $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$.
- Propiedad de centro de masa: $\sum_{i=1}^n (x_i - \bar{x}) = 0$.
- Cuando n es grande aproxima al verdadero valor de la media poblacional μ .
- La desventaja es que es muy sensible a la presencia de datos extremos en la muestra:
 - Muestra 1 = 37, 40, 46, 50, 57 $\Rightarrow \bar{x}_1 = 46$
 - Muestra 2 = 37, 40, 46, 57, **200** $\Rightarrow \bar{x}_2 = 76$
- La media no necesariamente tiene que estar en la muestra.

■ Mediana muestral

- Es el valor que deja el 50 % de las observaciones tanto por encima como por debajo de él.
- Es el valor central o el promedio de los dos valores centrales si n es impar o par respectivamente.
- $$\tilde{x} = \begin{cases} \frac{x_{\frac{n}{2}} + x_{(\frac{n}{2})+1}}{2} & \text{si } n \text{ es par} \\ x_{\frac{n+1}{2}} & \text{si } n \text{ es impar} \end{cases}$$
- Puede o no ser un valor de la muestra.

■ Percentil

- El percentil $i\%$ es aquel valor que acumula a su izquierda el $i\%$ de los datos.
- Notación: $(p(i))$.
- Cuando $i = 50 \Rightarrow p(50) = \tilde{x}$ es la mediana.

■ Cuartil Inferior (Q_1)

$$P(25) = \begin{cases} \text{Mediana de las } (\frac{n}{2}) \text{ observaciones} + \text{pequeñas si } n \text{ es par} \\ \text{Mediana de las } (\frac{n+1}{2}) \text{ observaciones} + \text{pequeñas si } n \text{ es impar} \end{cases}$$

■ Cuartil Superior (Q_3)

$$P(75) = \begin{cases} \text{Mediana de las } (\frac{n}{2}) \text{ observaciones} + \text{grandes si } n \text{ es par} \\ \text{Mediana de las } (\frac{n+1}{2}) \text{ observaciones} + \text{grandes si } n \text{ es impar} \end{cases}$$

1.6. Medidas de dispersión o variabilidad

- Rango
 - Es la diferencia entre el máximo valor de la muestra y el mínimo.
 - Es fácil de calcular y está en la misma medida que los datos muestrales.
 - Desventaja: considera sólo dos datos de la muestra:
 - Muestra 1 = 0, 5, 5, 5, 10.
 - Muestra 2 = 0, 4, 5, 6, 10.

Conclusión: la segunda muestra es más variable que la segunda pero el rango es el mismo (10).
- Rango Intercuartil
 - Es la diferencia entre el cuartil superior y el cuartil inferior.
 - $RIC = Q_3 - Q_1$.
- Varianza muestral
 - $S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$.
 - $S^2 \xrightarrow{n \rightarrow \infty} \sigma^2$ (Varianza poblacional).
 - Sensible a la presencia de datos extremos en la muestra ya que utiliza la media muestral:
 - Muestra A 100 valores iguales a 10.
 - Muestra B 99 valores iguales a 10 y uno igual a 1010.
 - $10 = S_A^2 = 0$ y $S_B^2 = 10000$
- Desviación estándar muestral
 - $S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$.
 - Sensible a la presencia de datos extremos en la muestra ya que utiliza la media muestral.
- Coeficiente de variación
 - $CV = \frac{S}{\bar{x}} \cdot 100 \%$.

- Permite comparar la variabilidad de características medidas en distintas escalas, luego la que tenga menor CV será el de menor variabilidad.
- El CV es adimensional.
- Ejemplo: medidas de alturas
 - Personas: $\bar{x} = 1,70m$ $S = 0,02m$ $CV = 1,18\%$
 - Edificios: $\bar{x} = 20m$ $S = 0,1m$ $CV = 0,50\%$

Luego el conjunto que tiene mayor variabilidad es de las alturas de personas.

1.7. Guía para la construcción de un gráfico de caja o box-plot

En 1977, Tukey presentó un simple método gráfico-cuantitativo que resume varias de las características más destacadas de un conjunto de datos.

Tal método se conoce con el nombre de gráfico de caja o **box-plot**.

Las características de los datos incorporadas por este gráfico son:

- Centro o posición del valor más representativo.
- Dispersión.
- Naturaleza y magnitud de cualquier desviación de la simetría.
- Identificación de los puntos no usuales o atípicos, es decir puntos marcadamente alejados de la masa principal de datos.

La presencia de datos atípicos producen cambios drásticos en la media muestral (\bar{x}) y la desviación estándar muestral (S), no así en otras medidas que son más resistentes o robustas, como lo son la mediana muestral \tilde{x} y el rango intercuartil (RIC).

Pasos a seguir para la construcción del box plot

- Paso 1: Ordenar los datos obtenidos en la muestra de menor a mayor.
- Paso 2: Calcular:
 - La media. (\bar{X})
 - La mediana muestral. (\tilde{X})
 - El cuartil inferior (Q_1) .
 - El cuartil superior (Q_3) .
 - El rango intercuartil. (RIC)
- Paso 3: Sobre un eje horizontal dibujar una caja cuyo borde izquierdo sea el cuartil inferior y el borde derecho el cuartil superior.
- Paso 4: Dentro de la caja marcar con un punto la posición del promedio muestral y trazar un segmento perpendicular cuya posición corresponde al valor de la mediana.
- Paso 5: Trazar segmentos desde cada extremo de la caja hasta las observaciones más alejadas, que no superen $(1, 5.RIC)$ de los bordes correspondientes.
- Paso 6: Si existen observaciones que superen $(1, 5.RIC)$ entonces marcarlos con circunferencias aquellos puntos comprendidos entre $(1, 5.RIC)$ y $(3.RIC)$ respecto del borde más cercano, estos puntos se llaman puntos **anómalos suaves**, y con asteriscos aquellos puntos que superen los $(3.RIC)$ respecto de los bordes más cercanos, estos puntos se llaman puntos **anómalos extremos**.

Cálculos necesarios para realizar el Gráfico de Caja para el Ejemplo 1.

$$\begin{aligned} \bar{x} &: 1,366667 & \tilde{x} &: 1,100000 \\ Q_1 &: 0,700000 & Q_3 &: 1,800000 \\ (RIC) &= Q_3 - Q_1 = 1,8 - 0,7 = 1,1 \end{aligned}$$

$$\begin{aligned} 1,5.(RIC) &= 1,65 & 3.(RIC) &= 3,3 \\ Q_1 - 1,5.(RIC) &= -0,95 & Q_1 - 3.(RIC) &= -2,6 \\ Q_3 + 1,5.(RIC) &= 3,45 & Q_3 + 3.(RIC) &= 5,1 \end{aligned}$$

Luego como el mínimo es 0.2, no hay datos atípicos en el extremo inferior.
Pero en el extremo superior hay tres observaciones que superan la distancia $(1,5.(RIC))$ respecto de Q_3 , ellos son: 3.6, 4.5 y 5.2.

Siendo los dos primeros atípicos suaves y el último atípico extremo.

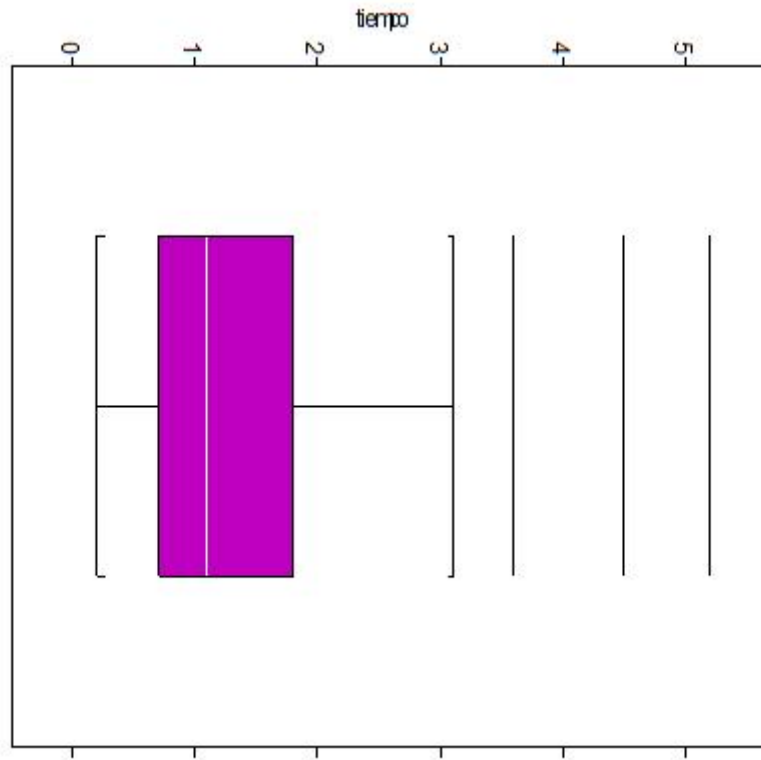


Figura 2: Gráfico de cajas del ejemplo de las cajas

Ejemplo

Los siguientes valores reflejan el contenido de un metabolito en la sangre de un paciente en 13 extracciones diferentes:

11,6	39,2	4,9	7,3	50,6	9,8	11,6
15,9	6,7	42,1	14,4	5,1	48,8	

Los datos están informados en $\frac{mg}{L}$.

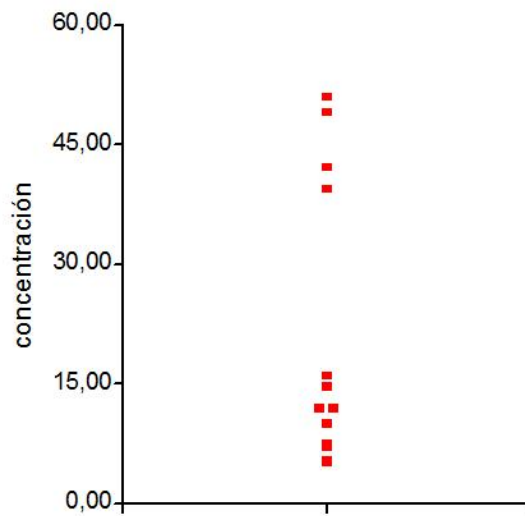


Figura 3: Gráfico de densidad de puntos de las extracciones de sangre

Ejemplo

Los datos que mostrare corresponden a una tesina de alumnas de la Escuela de Nutrición (Facultad de Medicina, UNC).

Tema de la tesina: ingesta de líquidos en el adulto mayor (AM).

Selección de la muestra: la muestra fue tomada de un grupo de AM que asisten al comedor del centro de jubilados de un barrio de la ciudad de Córdoba.

Algunos de los objetivos de este trabajo fueron:

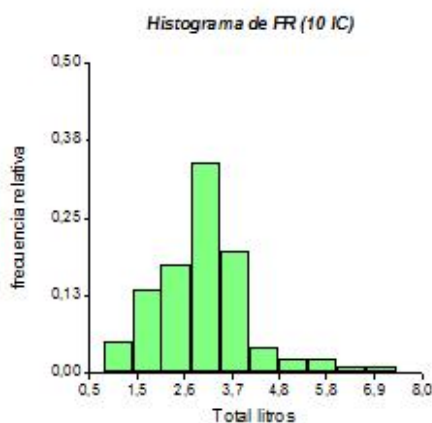
- Conocer la ingesta diaria de líquidos en AM, a partir de alimentos ricos en agua.
- Comparar la ingesta diaria de líquidos en AM por sexo.
- Determinar si los AM cumplen con las recomendaciones para la ingesta diaria de líquidos por sexo.

Las recomendaciones diarias de líquido por sexo son las siguientes: en mujeres debe ser de por lo menos 2,7 litros y en varones de por lo menos 3,7 litros.

Dos de las variables que consideraron son la ingesta diaria total de líquido (llamada Total litros) y el sexo del AM.

Estadística descriptiva para la variable ingesta diaria total de líquido

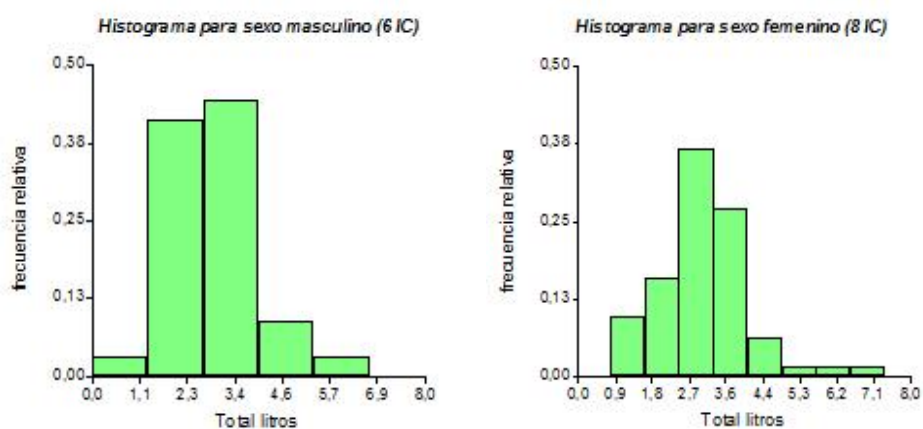
n	\bar{x}	S	S^2	Mín	Máx	\tilde{x}	Q_1	Q_3
97	3,0213	1,0920	1,1925	0,7888	7,3943	3,0495	2,3570	3,5496



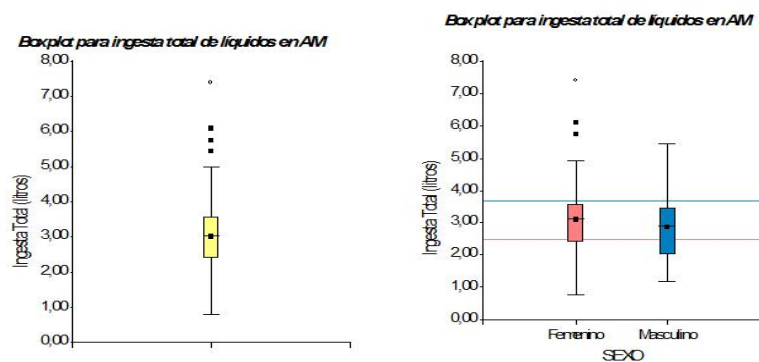
Estadística descriptiva para la variable ingesta diaria de líquido por sexo

SEXO	n	\bar{x}	S	S^2	Mín	Máx	\tilde{x}	Q_1	Q_3
F	63	3,0993	1,1344	1,2870	0,7888	7,3943	3,1168	2,3727	3,5877
M	34	2,8766	1,0089	1,0179	1,1947	5,4458	2,8861	1,9777	3,4463

Histogramas respectivos:



Gráficos de caja de los resultados obtenidos:



2. Probabilidad

Antes de poder hablar de probabilidad necesitamos definir los siguientes conceptos:

- **Espacio muestral** (“S”)

Definición: el espacio muestral es el conjunto de todos los valores posibles de un experimento

- **Evento**

Definición: llamaremos evento a todo subconjunto del espacio muestral, diremos que es simple si tiene un sólo elemento y compuesto si tiene más de un elemento.

- **Familia de eventos** (\mathcal{A})

Debe cumplir:

- $S \in \mathcal{A}$
- $A \in \mathcal{A} \Rightarrow \overline{A}^3 \in \mathcal{A}$
- $\{A_i\}_{i=1}^{\infty} : A_i \in \mathcal{A} \forall i \in \mathbb{N} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$

“La función de probabilidad”

Es una función $P: \mathcal{A} \rightarrow [0, 1]$ que satisface las siguientes condiciones:

1. $0 \leq P(A) \leq 1 \forall A \in \mathcal{A}$
2. $P(S) = 1$
3. $\{A_i\}_{i=1}^{\infty} : A_i \in \mathcal{A}$ y sean eventos disjuntos, id est: $(A_i \cap A_j = \emptyset \forall i \neq j)$

Entonces $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ donde A es un evento en \mathcal{A} .

“Modelo probabilístico” (S, \mathcal{A}, p)

Definición: llamaremos modelo probabilístico a la terna compuesta por el espacio muestral, la familia de eventos y una función de probabilidad.

Observación: sea S un espacio muestral tal que $\#S = n$ y cada punto de S es equiprobable entonces:

³Definimos \overline{A} como el complemento de A en S .

1. $P(E_i) = \frac{1}{n}$ para cada E_i evento simple en S.
2. Sea $A \in \mathcal{A} \Rightarrow P(A) = \frac{\#A}{\#S}$

Demostración

Sea (S, \mathcal{A}, p) un modelo probabilístico luego:

1. $S = \bigcup_{j=1}^n E_j : E_j$ es un evento simple y disjunto en S.

$$1 = P(S) = \sum_{j=1}^n \underbrace{P(E_j)}_{=p} = n \cdot p \therefore P(E_j) = \frac{1}{n}$$

2. Sea $A \in \mathcal{A} \Rightarrow A = \bigcup_{j \in J} E_j$ unión disjunta.

$$\text{Luego } P(A) = \sum_{j \in J} P(E_j) = \frac{j}{n} : j = \#A$$

■

Ejemplo

Dos estaciones de servicio tienen seis (6) bombas en uso, en cada estación para un tiempo fijo.

$$S = \{(i, j) : i, j \in \mathbb{Z} : 0 \leq i, j \leq 6\}$$

$$\#S = 49$$

$$\mathcal{A} = P(S)^4$$

Ejemplo

Consideremos el siguiente experimento: examinar un conjunto de baterías hasta encontrar a la primera que cumpla con las condiciones de calidad.

Denotamos con (E) si la batería cumple con las condiciones de calidad y con (F) en caso contrario.

Queda definido el espacio muestral

$$S = \left\{ E, FE, FFE, \dots, \overbrace{F \dots F}^j E, \dots \right\}$$

$$\text{Luego definimos } E_j = \underbrace{F \dots F}_j E : j \in \mathbb{Z}^{\geq 0} \therefore E_0 = E$$

$$\text{Luego } S = \bigcup_{j=0}^{\infty} E_j \text{ unión disjunta.}$$

⁴P(S) escualquier conjunto de S.

Sea p la probabilidad de que una batería cumpla con la condición de calidad $\Rightarrow (1 - p)$ es la probabilidad de F.

$$P(E_j) = (1 - p)^j \cdot p : j \in \mathbb{Z}^{\geq 0}$$

$$P(S) = \sum_{j=0}^{\infty} P(E_j) = p \cdot \sum_{j=0}^{\infty} (1 - p)^j \stackrel{*}{=} p \cdot \frac{1}{1 - (1 - p)} = 1$$

$$\star \sum_{j=0}^{\infty} q^j = \frac{1}{1 - q} : 0 < q < 1$$

$$A \in \mathcal{A} \Rightarrow A = \bigcup_{j \in J} E_j \therefore P(A) = \sum_{j \in J} P(E_j)$$

■

2.1. Propiedades de un modelo probabilístico

Sea (S, \mathcal{A}, p) un modelo probabilístico entonces vale que:

1. $A \in \mathcal{A} \Rightarrow P(\overline{A}) = 1 - P(A)$
2. Sean $A, B \in \mathcal{A} : A \subset B \Rightarrow P(B - A) = P(B) - P(A)$ y $P(B) \geq P(A)$
3. Sean $A, B \in \mathcal{A} \Rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Demostración

1. Sean $A, B \in \mathcal{A} \Rightarrow S = A \cup \overline{A}$ unión disjunta.

2. Podemos escribir a B como una unión disjunta de la siguiente manera:

$$B = A \cup (B - A)$$

S

Luego la $P(B) \stackrel{Disj}{=} P(A) + P(B - A)$

$\therefore P(B - A) = P(B) - P(A)$ y $P(B - A) \geq 0 \Rightarrow P(B) \geq P(A)$

3. $A \cup B = A \cup (B - (A \cap B))$ unión disjunta.

$$\begin{aligned}
P(A \cup B) &= P(A \cup (B - (A \cap B))) \\
&= P(A) + P(B - (A \cap B)) \\
&= P(A) + P(B) - P(A \cap B)
\end{aligned} \tag{1}$$

$$(1) \text{ Como } (A \cap B) \subset B \Rightarrow P(B - (A \cap B)) \stackrel{2}{=} P(B) - P(A \cap B)$$

Ejemplo

En un lugar desconocido el 60 % de las familias compran el diario A y el 80 % de las familias compran el diario B. Por otra parte sabemos que el 50 % está suscripta a ambos diarios.

Se pide calcular:

1. ¿Cuál es la probabilidad de tomar una familia que compre algún diario?

$$\begin{aligned}
P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\
&= 0,6 + 0,8 - 0,5 \\
&= 0,9
\end{aligned}$$

2. ¿Cuál es la probabilidad de tomar una familia en el conjunto que compre sólo un diario?

$$\begin{aligned}
P(A - (A \cap B) \cup (B - (A \cap B))) &= P(A - (A \cap B)) + P(B - (A \cap B)) \\
&= P(A) - P(A \cap B) + P(B) - P(A \cap B) \\
&= 0,6 - 0,5 + 0,8 - 0,5 \\
&= 0,4
\end{aligned}$$

2.2. Probabilidad condicional

Definición: sea (S, \mathcal{A}, p) un modelo probabilístico, con $B \in \mathcal{A}$ tal que $P(B) > 0$, entonces llamamos probabilidad de A dado B como:

$$P_B(A) \stackrel{Def}{=} P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \forall A \in \mathcal{A}$$

2.2.1. Propiedades de la probabilidad condicional

1. $0 \leq P(A|B) \leq 1 \forall A \in \mathcal{A}$
2. $P_B(S) = P(S|B) = 1$
3. Sea $\{A_i\}_{i=1}^{\infty}$ una sucesión de eventos disjuntos $\in \mathcal{A}$ entonces:

$$P_B\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P_B(A_i)$$

Demostración:

1. Sea $A \in \mathcal{A} \Rightarrow 0 \leq P(A|B) = \frac{P(A \cap B)}{P(B)} \leq^* 1$
 $* \text{ Ya que } (A \cap B) \subset B \Rightarrow P(A \cap B) \leq P(B)$
2. $P_B(S) = P(S|B) \stackrel{Def}{=} \frac{P(S \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$
3. Quiero ver que: $P_B\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P_B(A_i)$

$$\begin{aligned} P_B\left(\bigcup_{i=1}^{\infty} A_i\right) &= P\left(\bigcup_{i=1}^{\infty} A_i | B\right) \\ &= \frac{P\left(\bigcup_{i=1}^{\infty} A_i \cap B\right)}{P(B)} \\ &= \frac{\sum_{i=1}^{\infty} P(A_i \cap B)}{P(B)} \\ &= \sum_{i=1}^{\infty} P_B(A_i) \end{aligned} \tag{2}$$

(2) Por ser disjuntos.

La regla de la multiplicación

Sea (S, \mathcal{A}, p) un modelo probabilístico, y sean $\{A_i\}_{i=1}^n$ eventos en \mathcal{A} .
 Entonces:

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \dots (P(A_n | \bigcap_{i=1}^{n-1} A_i))$$

Ley de probabilidad total

Sea A_1, \dots, A_n eventos disjuntos en $\mathcal{A} : S = \bigcup_{i=1}^n A_i$ y sea $B \in \mathcal{A}$ tenemos que:

$$P(B) = \sum_{i=1}^n P(A_i) \cdot P(B|A_i)$$

2.3. Teorema de Bayes

Sea A_1, \dots, A_n eventos disjuntos en $\mathcal{A} : S = \bigcup_{i=1}^n A_i$ y sea $B \in \mathcal{A}$ entonces:

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i) \cdot P(B|A_i)}{\sum_{j=1}^n P(A_j) \cdot P(B|A_j)}$$

Ejemplo

Existe un negocio que vende 3 marcas de dvds: 1, 2, 3, que son elegidos por los clientes asiduamente con un porcentaje del 50 %, 30 %, 20 %, respectivamente, con un año de garantía.

Antes de que termine la garantía los dvds de la marca 1 terminan rompiéndose en un 50 % de los casos, los de la marca 2 en un 20 % y los de la marca 3 en un 10 %.

Definimos $A_i =$ “el cliente compra la marca i” : $1 \leq i \leq 3$

Se pide:

1. Calcular la probabilidad de que si se compra un dvd, en este negocio, deba ser reparado antes del año de garantía.

$$\begin{aligned} P(R) &= P(A_1 \cap R) + P(A_2 \cap R) + P(A_3 \cap R) = \sum_{i=1}^3 P(A_i) \cdot P(R|A_i) \\ &= P(A_1) \cdot P(R|A_1) + P(A_2) \cdot P(R|A_2) + P(A_3) \cdot P(R|A_3) \\ &= 0,5 * 0,25 + 0,3 * 0,2 + 0,2 * 0,1 \\ &= 0,205 \end{aligned}$$

2. Dado que el dvd requiere una reparación, ¿cuál es la probabilidad de que haya comprado la marca “i”?

$$P(A_i|R) = \frac{P(A_i).P(R|A_i)}{P(R)} : i \in \{1, 2, 3\}$$

Y se obtiene:

$$P(A_1|R) \approx 0,61 \quad P(A_2|R) \approx 0,29 \quad P(A_3|R) \approx 0,1$$

2.4. Eventos independientes

Definición: diremos que dos eventos son independientes si:

$$P(A \cap B) = P(A).P(B)$$

Proposición: Si A y B son eventos en \mathcal{A} se tiene que las siguientes son equivalentes:

- A y B son independientes
- A y \overline{B} son independientes
- \overline{A} y B son independientes
- \overline{A} y \overline{B} son independientes

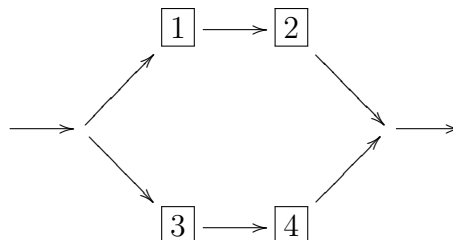
Definición: Sean A_1, \dots, A_N eventos en \mathcal{A} .

Diremos que son mutuamente excluyentes si:

$$P\left(\bigcap_{i=1}^r A_{i, k}\right) = \prod_{k=1}^r P(A_{i, k}) \quad \forall (i, k)_{k=1}^r \subset \{1, \dots, n\}$$

Ejemplo

Supongamos 4 componentes conectados de la siguiente manera:



El sistema funciona (F) si funcionan las componentes 1 y 2 o si funcionan las componentes 3 y 4.

Supongamos que las componentes funcionan independientemente una de otra y que además la probabilidad de que funcione cada componente es del 90 %.

Se pide calcular la probabilidad de que funcione el sistema.

Sea A_i = “funciona el componente i-ésimo”: $i=1, 2, 3, 4$.

$$F = (A_1 \cap A_2) \cup (A_3 \cap A_4)$$

$$\begin{aligned} P(F) &= P((A_1 \cap A_2) \cup (A_3 \cap A_4)) \\ &= P(A_1 \cap A_2) + P(A_3 \cap A_4) - P((A_1 \cap A_2) \cap (A_3 \cap A_4)) \\ &= (0,9)^2 + (0,9)^2 - (0,9)^4 \\ &= 0,9639 \end{aligned}$$

2.5. Variables aleatorias

Definición: dado (S, \mathcal{A}, p) un modelo probabilístico llamaremos variable aleatoria (V.A.) a cualquier función:

$$X : S \rightarrow \mathbb{R} : [X \leq x] = \{w \in S : X(w) \leq x\} \in \mathcal{A} \forall x \in \mathbb{R}$$

2.5.1. Clasificación de las variables aleatorias

- Diremos que X es una variable aleatoria discreta si toma un conjunto finito o infinito numerable de valores posibles con probabilidad positiva.
- Diremos que X es una variable aleatoria Bernoulli si toma sólo dos valores posibles: 1 ó 0.
- Diremos que X es una variable aleatoria continua si sus valores posibles consisten en un todo intervalo en la recta numérica.

3. Variables aleatorias discretas

Sea X una variable aleatoria discreta definimos la función de distribución de probabilidad o función de probabilidad de masa como la función:

$$P : \mathbb{R} \rightarrow [0, 1] : P(x) = P[X = x]$$

Llamaremos función de distribución acumulada de X a la función

$$F : \mathbb{R} \rightarrow [0, 1] \text{ dada por } F(x) = P(X \leq x) \quad \forall x \in \mathbb{R}$$

Ejemplo

Si quisiéramos definir la función de distribución de probabilidad de X , donde X cuenta la cantidad de caras que se obtienen al tirar 5 veces una moneda honesta.

$$S = \{\overbrace{(x_1, x_2, x_3, x_4, x_5)}^w : x_i \text{ es } 1(\text{cara}) \text{ ó } 0(\text{cruz})\}$$

$$\mathcal{A} = P(s) \Rightarrow P(w) = \frac{1}{32} \quad \forall w \in S$$

$$[X \leq x] = \{w \in S : X(w) \leq x\}$$

$$\text{Si } x < 0 \Rightarrow [X \leq x] = \emptyset \in \mathcal{A}$$

$$\text{Si } x \geq 5 \Rightarrow [X \leq x] = S \in \mathcal{A}$$

$$\text{Si } x \in [0, 5] \Rightarrow [X \leq x] = \{\bigcup (x_1, x_2, x_3, x_4, x_5) : \sum_{i=1}^5 x_i \leq x\} \in \mathcal{A}$$

\therefore Es una variable aleatoria de tipo discreta ya que toma un número finito de valores.

$$\text{Si } x \notin \{0, 1, 2, 3, 4, 5\} \Rightarrow P(x) = P(X = x) = P(\emptyset) = 0$$

$$P(0) = P(X = 0) = \frac{1}{32}$$

$$P(1) = P(X = 1) = \{(10000), (01000), (00100), (000010), (00001)\} = \frac{5}{32}$$

$$P(2) = P(X = 2) = \frac{1}{32} \binom{5}{2} = \frac{10}{32}$$

$$P(3) = P(X = 2) = \frac{1}{32} \binom{5}{3} = \frac{10}{32}$$

$$P(4) = P(X = 2) = \frac{1}{32} \binom{5}{4} = \frac{5}{32}$$

$$P(5) = P(X = 2) = \frac{1}{32} \binom{5}{5} = \frac{1}{32}$$

$$P(x) = \begin{cases} \frac{1}{5} \cdot \binom{5}{x} & \text{si } 0 \leq x \leq 5 \\ 0 & \text{en caso contrario} \end{cases}$$

Luego si quisiéramos hallar la función de distribución acumulada de X:

$$\text{Si } x < 0 \Rightarrow F(x) = P(X \leq x) = P(\emptyset) = 0$$

$$\text{Si } x \in [0, 1) \Rightarrow F(x) = P(X \leq 0) = P(X = 0) = \frac{1}{32}$$

$$\begin{aligned} \text{Si } x \in [1, 2) \Rightarrow F(x) &= P(X \leq 1) = P((X = 0) \cup (X = 1)) \\ &= P(0) + P(1) = \left(\binom{5}{0} + \binom{5}{1} \right) \frac{1}{32} = \frac{6}{32} \end{aligned}$$

$$\begin{aligned} \text{Si } x \in [2, 3) \Rightarrow F(x) &= P(X \leq 1) = P((X = 0) \cup (X = 1) \cup (X = 2)) \\ &= P(0) + P(1) + P(2) = \frac{16}{32} \end{aligned}$$

$$\begin{aligned} \text{Si } x \in [3, 4) \Rightarrow F(x) &= P(X \leq 1) = P((X = 0) \cup (X = 1) \cup (X = 2)) \\ &= P(0) + P(1) + P(2) + P(3) = \frac{26}{32} \end{aligned}$$

$$\begin{aligned} \text{Si } x \in [4, 5) \Rightarrow F(x) &= P(X \leq 1) = P((X = 0) \cup (X = 1) \cup (X = 2)) \\ &= P(0) + P(1) + P(2) + P(3) + P(4) = \frac{31}{32} \end{aligned}$$

$$\text{Si } x \geq 5 \Rightarrow F(x) = P(X \leq 5) = 1$$

3.1. Propiedades de una variable aleatoria discreta

- $\sum_X P(x) = 1$
- La función es monótona creciente, id est si $x_1 < x_2 \Rightarrow F(x_1) < F(x_2)$
- La gráfica de F es escalonada y tiene discontinuidades en los distintos valores que toma la variable aleatoria.
- $\lim_{x \rightarrow -\infty} F(x) = 0$
- $\lim_{x \rightarrow \infty} F(x) = 1$
- $P(x) = P(X = x) = F(x) - \lim_{t \rightarrow x^-} F(t) \quad \forall x \in \mathbb{R}$

Volviendo al ejemplo de las baterías

Primero veamos que $X(E_j) = j + 1 \ \forall j \in \mathbb{Z}^{\geq 0}$ es una variable aleatoria.

Si $x < 1 : [X \leq x] = \emptyset \in \mathcal{A}$

Si $x \geq 1 : [X \leq x] = \bigcup_{j \leq [x]-1} E_j \in \mathcal{A}.$

Observación: $[x]$ denota la parte entera de x .

$\therefore X$ es una variable aleatoria discreta

“Queremos hallar la función de probabilidad”

$$\text{Si } x \notin \mathbb{N} \Rightarrow P(x) = P(\underbrace{X = x}_{= \emptyset}) = 0$$

$$P(1) = P(X = 1) = P(E_0) = 0$$

$$P(2) = P(X = 2) = P(E_1) = p \cdot (1 - p)$$

\vdots

$$P(n + 1) = P(X = n + 1) = p \cdot (1 - p)^n = P(E_n) \ \forall n \in \mathbb{N}$$

“Luego queremos hallar la función de distribución acumulada de X ”

Si $x < 1 \Rightarrow [X \leq x] = \emptyset \Rightarrow F(x) = 0$

Sea $x \in [k, k + 1) : k \in \mathbb{N}$

$$\begin{aligned} F(x) &= \sum_{i=1}^k P(i) = \sum_{i=1}^{[x]} P(i) \\ &= p \cdot \sum_{i=1}^{[x]} (1 - p)^{i-1} \end{aligned} \tag{3}$$

$$= p \cdot \sum_{j=0}^{[x]+1} (1 - p)^j \tag{4}$$

$$\begin{aligned} &= p \cdot \frac{1 - (1 - p)^{[x]+1}}{1 - (1 - p)} \\ &= p * \frac{1 - (1 - p)^{[x]+1}}{p} \\ &= 1 - (1 - p)^{[x]+1} \end{aligned} \tag{5}$$

(3) Reemplazo $P(i)$.

(4) Hago un cambio de variables llamando a: $j = i + 1$.

(5) “Serie telescópica”: $\sum_{j=0}^{n-1} q^j = \frac{1-q^n}{1-q}$.

$$\therefore F(x) = \begin{cases} 1 - (1 - p)^x & \forall x \in \mathbb{N} \\ 0 & \text{en caso contrario} \end{cases}$$

Ejemplo

Sea una población M, donde los habitantes contraen dos tipos de enfermedades distintas: A y B.

Supongamos que sabemos que:

- El 30 % de la población no posee ninguna de las dos enfermedades.
- El 35 % de la población tiene la enfermedad A.
- El 55 % de la población tiene la enfermedad B.

Sea X la cantidad de enfermedades de un sujeto de la población M.

$$Im(x) = \{0, 1, 2\}$$

Se pide hallar la función de distribución de probabilidad de X:

$$\text{Si } x \notin Im(x) \Rightarrow P(x) = 0$$

$$\text{Si } x = 0 \Rightarrow P(0) = P(X = 0) = 0,30$$

$$\text{Si } x = 1 \Rightarrow P(1) = P(X = 1) = 0,50$$

$$\text{Si } x = 2 \Rightarrow P(2) = P(X = 2) = 0,20$$

$$\therefore P(x) = \begin{cases} 0,3 & \text{si } x = 0 \\ 0,5 & \text{si } x = 1 \\ 0,2 & \text{si } x = 2 \\ 0 & \text{en caso contrario} \end{cases}$$

Luego obtenemos la función de distribución acumulada: “F(x)”

$$\text{Si } x < 0 \Rightarrow F(0) = P(X < 0) = P(\emptyset) = 0$$

$$\text{Si } x \in [0, 1) \Rightarrow F(x) = P(X \leq 0) = P(X = 0) = 0,30$$

$$\text{si } x \in [1, 2) \Rightarrow F(x) = P(X \leq 1) = P(X = 0) + P(X = 1) = 0,80$$

$$\text{si } x \geq 2 \Rightarrow F(x) = P(S) = 1$$

$$\therefore F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 0,3 & \text{si } x \in [0, 1) \\ 0,8 & \text{si } x \in [1, 2) \\ 1 & x \geq 2 \end{cases}$$

3.2. Valor esperado de una variable aleatoria discreta

Definición: sea X una variable aleatoria discreta con función de distribución de probabilidad $P(x)$, si $\sum_{i \in I} |x_i| \cdot P(x_i) < \infty$. Con $Im(x) = \{x_i\}_{i \in I}$ llamaremos valor esperado o valor muestral o esperanza de X a:

$$E(x) = \sum_{i \in I} x_i \cdot P(x_i) = \mu$$

Observación: en el ejemplo anterior $E(x) = 0 \cdot 0,3 + 1 \cdot 0,5 + 5 \cdot 0,2 = 0,9$

Propiedad: sea X una variable aleatoria discreta con función de distribución de probabilidad $P(x)$ y sea $w = h(x)$ entonces si $\sum_{i \in I} h(x_i) \cdot P(x_i) < \infty$

Definimos: $E(h(x)) = \sum_{i \in I} h(x_i) \cdot P(x_i) : Im(x) = \{x_i\}_{i \in I}$

Ejemplo

$$Sea h(x) = x^2 \Rightarrow E(x^2) = \sum_{i \in I} x_i^2 \cdot P(x_i)$$

3.3. Varianza de una variable aleatoria discreta

Definición: sea X una variable aleatoria discreta con función de distribución de probabilidad $P(x)$ llamaremos varianza de X a:

$$\sigma^2 = V(x) = E(x - \mu)^2 : \mu = E(x)$$

Definición: sea X una variable aleatoria discreta con función de distribución de probabilidad $P(x)$ llamaremos **desvío estándar** de x a:

$$\sigma = \sqrt{V(x)}$$

3.4. Propiedades de la varianza y la esperanza de una V.A. discreta

- $E(a \cdot x + b) = a \cdot E(x) + b \quad \forall a, b \in \mathbb{R}$
- $V(x) \geq 0$ (Por definición)
- $V(x) = E(x^2) - \mu^2 = E(x^2) - E(x)^2$
- $V(a \cdot x + b) = a^2 \cdot V(x) \quad \forall a, b \in \mathbb{R}$

Demostración:

- Queremos ver que $E(a.x + b) = a.E(x) + b \forall a, b \in \mathbb{R}$

$$\begin{aligned}
 E(\underbrace{a.x + b}_{h(x)}) &= \sum_{i \in I} (a.x_i + b).P(x_i) \\
 &= a. \underbrace{\sum_{i \in I} x_i.P(x_i)}_{E(x)} + b. \underbrace{\sum_{i \in I} P(x_i)}_{= 1} \\
 &= a.E(x) + b
 \end{aligned}$$

- Queremos ver que $V(x) = E(x^2) - E(x)^2$

$$\begin{aligned}
 V(x) &= \sum_{i \in I} (x_i - \mu)^2.P(x_i) \\
 &= \sum_{i \in I} x_i^2.P(x_i) + \mu^2. \underbrace{\sum_{i \in I} P(x_i)}_{= 1} - 2.\mu. \underbrace{\sum_{i \in I} x_i.P(x_i)}_{= \mu} \\
 &= E(x^2) + \mu^2 - 2.\mu^2 \\
 &= E(x^2) - \mu^2
 \end{aligned}$$

- Queremos ver que $V(a.x + b) = a^2.V(x) \forall a, b \in \mathbb{R}$

$$\begin{aligned}
 V(a.x + b) &= E((a.x + b) - E(a.x + b))^2 \\
 &= E((a.x + b) - (a.\mu + b))^2 \\
 &= E(a^2.(x - \mu)^2) \\
 &= a^2.E(x - \mu)^2 \\
 &= a^2.V(x)
 \end{aligned}$$

Ejemplos

- Sea X una Bernoulli(p) tenemos que:

$$P(0) = P(X = 0) = 1 - p \qquad P(1) = P(X = 1) = p$$

$$E(x) = 0 \cdot (1 - p) + 1 \cdot p = p \qquad E(x^2) = 0^2 \cdot (1 - p) + 1^2 \cdot p = p$$

$$\therefore V(x) = E(x^2) - E(x)^2 = p - p^2 = p \cdot (1 - p) = p \cdot q$$

- Sea X = “número que se obtiene en la tirada de un dado honesto”

$$Im(x) = \{1, 2, 3, 4, 5, 6\}$$

$$P(i) = P(X = i) = \frac{1}{6} \quad \forall i \in Im(x)$$

$$E(x) = \sum_{i=1}^6 i \cdot P(i) = \frac{1}{6} \sum_{i=1}^6 i = \frac{1}{6} * \frac{6 * 7}{2} = 3,5 \qquad (6)$$

$$E(x^2) = \sum_{i=1}^6 i^2 \cdot P(i) = \frac{1}{6} \sum_{i=1}^6 i^2 = \frac{1}{6} * \frac{6 * 7 * 13}{6} = 15,16 \qquad (7)$$

$$(6) \quad \sum_{i=1}^n i = \frac{n \cdot (n+1)}{2}$$

$$(7) \quad \sum_{i=1}^n i^2 = \frac{n \cdot (n+1) \cdot (2n+1)}{6}$$

$$\therefore V(x) = E(x^2) - E(x)^2 = 15,16 - (3,5)^2 = 2,91$$

- Volvamos al ejemplo de las baterías

Definimos $X =$ “número de baterías que deben ser probadas hasta obtener la primera que cumpla las condiciones de calidad”

Luego la función de distribución de probabilidad nos queda:

$$P(x) = \begin{cases} p \cdot (1-p)^{x-1} & \forall x \in \mathbb{N} \\ 0 & \text{en caso contrario} \end{cases}$$

Se pide hallar su esperanza y su varianza:

$$\begin{aligned} E(x) &= \sum_{i=1}^{\infty} i \cdot P(i) = \sum_{i=1}^{\infty} i \cdot p \cdot (1-p)^{i-1} \\ &= p \cdot \sum_{i=1}^{\infty} i \cdot (1-p)^{i-1} \\ &= p \cdot \frac{d}{dp} \cdot (-1) \cdot \left(\left(\sum_{i=1}^{\infty} (1-p)^i + 1 \right) - 1 \right) \\ &= p \cdot \frac{d}{dp} \cdot (-1) \cdot \left(\sum_{i=0}^{\infty} (1-p)^i - 1 \right) \\ &= p \cdot \frac{d}{dp} \cdot (-1) \cdot \left(\frac{1}{1 - (1-p)} - 1 \right) \\ &= p \cdot \frac{d}{dp} \cdot (-1) \cdot \left(\frac{1}{1-p} - 1 \right) \\ &= (-1) \cdot p \cdot \frac{-1}{p^2} \\ &= \frac{1}{p} \end{aligned}$$

Para calcular la varianza necesitamos saber la $E(x^2)$:

$$\begin{aligned}
 E(x^2) &= \sum_{i=0}^{\infty} i^2 \cdot P(i) = \sum_{i=1}^{\infty} i^2 \cdot P(i) \\
 &= \sum_{i=2}^{\infty} i \cdot (i-1) \cdot P(i) + \underbrace{\sum_{i=0}^{\infty} i \cdot P(i)}_{E(x)} \\
 &= \frac{2 \cdot (1-p)}{p^2} + \frac{1}{p} \\
 &= \frac{2-p}{p^2}
 \end{aligned} \tag{8}$$

$$\begin{aligned}
 (8) \quad \sum_{i=2}^{\infty} i \cdot (i-1) \cdot P(i) &= p \cdot (1-p) \cdot \sum_{i=2}^{\infty} i \cdot (i-1) (1-p)^{i-2} \\
 &= p \cdot (1-p) \cdot \frac{d^2}{dp^2} \left(\sum_{i=2}^{\infty} (i-p)^i \right) \\
 &= p \cdot (1-p) \cdot \frac{d^2}{dp^2} \left(\sum_{i=0}^{\infty} (i-p)^i - (1-p) - 1 \right) \\
 &= p \cdot (1-p) \cdot \frac{d^2}{dp^2} \left(\frac{1}{p} - 2 + p \right) \\
 &= p \cdot (1-p) \cdot \frac{d}{dp} \left(\frac{-1}{p^2} + 1 \right) \\
 &= p \cdot (1-p) \cdot \left(\frac{2}{p^3} \right) \\
 &= \frac{2 \cdot (1-p)}{p^2}
 \end{aligned}$$

$$\therefore V(x) = E(x^2) - E(x)^2 = \frac{1-p}{p^2}$$

- De una urna con n cartones numerados del 1 al n se extrae uno al azar. Sea X el número obtenido, se pide hallar la esperanza de X y su varianza.

$$Im(X) = \{i \in \mathbb{N} : 1 \leq i \leq n\}$$

$$P(i) = \frac{1}{n} \quad \forall i \in Im(x)$$

$$E(x) = \sum_{i=1}^n i \cdot P(i) = \frac{1}{n} \cdot \sum_{i=1}^n i = \frac{1}{n} \cdot \frac{n \cdot (n+1)}{2} = \frac{n+1}{2}$$

$$E(x^2) = \sum_{i=1}^n i^2 \cdot P(i) = \frac{1}{n} \cdot \sum_{i=1}^n i^2 = \frac{1}{n} \cdot \frac{n \cdot (n+1) \cdot (2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$$

$$V(x) = E(x^2) - E(x)^2 = \frac{(n+1)(2n+1)}{6} - \frac{n+1}{2} = \frac{n^2-1}{12}$$

3.5. Distribución Binomial

Definición: diremos que un experimento es binomial si cumple las siguientes condiciones:

- Consta de n ensayos idénticos.
- Cada ensayo tiene sólo dos resultados posibles E = éxito ó F = fracaso.
- Los ensayos son independientes uno de otro.
- La $P(E) = p$ para cada uno de los ensayos.

Definición: sea X la variable que cuenta el número de éxitos en un experimento binomial. Entonces diremos que X tiene “distribución binomial de parámetros n y p ”.

Notación: $X \sim B(n, p)$

Ejemplo

Sea un experimento donde tiramos 5 veces una moneda honesta. Definimos X como la variable aleatoria que cuenta el número de caras (E) en las 5 tiradas.

$$P(E) = 0,5 \therefore X \sim B(5, 0,5)$$

3.5.1. Función de distribución de probabilidad de una binomial

Dado un experimento binomial, sea X la variable aleatoria que cuenta los éxitos (E) en n ensayos independientes con $P(E) = p$ para cada ensayo.

$$S = \{(x_1, \dots, x_n) : x_i = E \vee x_i = F \forall i : 1 \leq i \leq n\}$$

$$Im(x) = \{k \in \mathbb{Z} : 0 \leq k \leq n\}$$

Tenemos que: $P(k) = P(X = k)$

Si $k \notin Im(x) \Rightarrow P(k) = P(\emptyset) = 0$

Si $k \in Im(x) \Rightarrow [X = k] = \{x_1, \dots, x_k, \dots, x_n\} \in S$ tal que hay k éxitos y $(n-k)$ fracasos

$$\#[X = k] = \binom{n}{k}$$

$$Si P(E) = p \Rightarrow P(F) = 1 - p$$

$$P\left(\underbrace{\{E, \dots, E\}}_k, \underbrace{\{F, \dots, F\}}_{n-k}\right) \stackrel{*}{=} p^k \cdot (1-p)^{n-k}$$

★ Por ser independientes y por ser p constante en cada uno de los ensayos.

$$\therefore P(k) = P(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \forall k \in Im(x)$$

Luego queremos hallar la esperanza y la varianza:

3.5.2. Esperanza de una binomial

$$\begin{aligned}
 E(x) &= \sum_{k=0}^n k \cdot P(k) \\
 &= \sum_{k=0}^n k \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}
 \end{aligned} \tag{9}$$

$$= \sum_{k=1}^n k \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \tag{10}$$

$$= n \cdot \sum_{k=1}^n \binom{n-1}{k-1} \cdot p^k \cdot (1-p)^{n-k} \tag{11}$$

$$\begin{aligned}
 &= n \cdot p \cdot \sum_{k=1}^n \binom{n-1}{k-1} \cdot p^{k-1} \cdot (1-p)^{(n-1)-(k-1)} \\
 &= n \cdot p \cdot \sum_{j=0}^{n-1} \binom{n-1}{j} \cdot p^j \cdot (1-p)^{n-1-j}
 \end{aligned} \tag{12}$$

$$\begin{aligned}
 &= n \cdot p \cdot (p + (1-p))^{n-1} \\
 &= n \cdot p
 \end{aligned} \tag{13}$$

(9) Definición de $P(k)$.

(10) $k = 0 \Rightarrow \Sigma = 0$.

(11) $k \cdot \binom{n}{k} = \frac{k \cdot n!}{k!(n-k)!} = \frac{n \cdot (n-1)!}{(k-1)!((n-1)-(k-1))!} = n \cdot \binom{n-1}{k-1}$.

(12) Hacemos un cambio de variable llamando a: $j = k + 1$.

(13) Por binomio de newton $(a + b)^m = \sum_{r=0}^m \binom{m}{r} \cdot a^r \cdot b^{m-r}$.

Para calcular $V(x)$ primero despejaremos la $E(x^2)$

$$E(x^2) = \sum_{k=0}^n k^2 \cdot P(k)$$

$$= \sum_{k=0}^n k^2 \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \quad (14)$$

$$= \sum_{k=1}^n k^2 \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \quad (15)$$

$$= \sum_{k=1}^n k \cdot (k-1) \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} + \underbrace{\sum_{k=1}^n k \cdot P(k)}_{E(x)=n \cdot p} \quad (16)$$

$$= \sum_{k=1}^n k \cdot (k-1) \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} + n \cdot p$$

$$= \sum_{k=2}^n k \cdot (k-1) \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} + n \cdot p \quad (17)$$

$$= \sum_{k=2}^n k \cdot (k-1) \cdot \frac{n!}{k!(n-k)!} \cdot p^k \cdot (1-p)^{n-k} + n \cdot p$$

$$= n \cdot (n-1) \sum_{k=2}^n \frac{(n-2)!}{(k-2)!((n-2)-(k-2))!} \cdot p^k \cdot (1-p)^{n-k} + n \cdot p$$

$$= n \cdot (n-1) \cdot p^2 \sum_{k=2}^n \binom{n-2}{k-2} \cdot p^{k-2} \cdot (1-p)^{n-k} + n \cdot p$$

$$= n \cdot (n-1) \cdot p^2 \underbrace{\sum_{j=0}^{n-2} \binom{n-2}{j} \cdot p^j \cdot (1-p)^{n-2-j}}_{\Sigma=1} + n \cdot p \quad (18)$$

$$= n \cdot p \cdot ((n-1) \cdot p + 1)$$

(14) Definición de $P(k)$.

(15) $k = 0 \Rightarrow \Sigma = 0$.

(16) Reescribimos $k^2 = (k \cdot ((k-1) + 1))$.

(17) $k = 1 \Rightarrow \Sigma = 0$.

(18) Hago un cambio de variables llamando $j = k - 2$.

Finalmente puedo calcular la varianza de una binomial:

$$\begin{aligned}
 \therefore V(x) &= E(x^2) - E(x)^2 \\
 &= n.p.((n-1).p + 1 - n.p) \\
 &= n.p.(1-p) \\
 &= n.p.q = n.P(E).P(F)
 \end{aligned}$$

En síntesis sea $X \sim B(n, p)$

- $P(k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \quad \forall k \in Im(x)$
- $E(x) = n.p$
- $V(x) = n.p.(1-p) = n.p.q = n.P(E).P(F)$

3.6. Distribución de Poisson

Definición: diremos que una variable aleatoria X tiene distribución de Poisson de parámetros $\lambda : \lambda > 0$ si su función de probabilidad está dada por

$$P(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!} \text{ con } k \in Im(x) = \mathbb{Z}^{\geq 0}$$

Notación: $X \sim P(\lambda)$

$$e^{\lambda} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \tag{19}$$

$$1 = \sum_{k=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{k!} = \sum_{k=0}^{\infty} P(k) \tag{20}$$

(19) Por el desarrollo de Taylor.

(20) Multiplico por $e^{-\lambda}$.

3.6.1. Esperanza y varianza de una Poisson

$$\begin{aligned}
 E(x) &= \sum_{k=0}^{\infty} k.P(k) \\
 &= \sum_{k=1}^{\infty} k.P(k) \\
 &= \sum_{k=1}^{\infty} k.e^{-\lambda} \cdot \frac{\lambda^k}{k!} \\
 &= \sum_{k=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k}{(k-1)!} \\
 &= \sum_{k=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda \cdot \lambda^{k-1}}{(k-1)!} \\
 &= e^{-\lambda} \cdot \lambda \cdot \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
 &= e^{-\lambda} \cdot \lambda \cdot \underbrace{\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}}_{e^{\lambda}} \\
 &= \lambda
 \end{aligned} \tag{21}$$

$$(21) \quad k = 0 \Rightarrow \Sigma = 0$$

$$\begin{aligned}
E(x^2) &= \sum_{k=0}^{\infty} k^2 \cdot P(k) \\
&= \sum_{k=1}^{\infty} k^2 \cdot P(k)
\end{aligned} \tag{22}$$

$$= \sum_{k=1}^n k \cdot (k-1) \cdot P(k) + \underbrace{\sum_{k=1}^n k \cdot P(k)}_{E(x)=\lambda} \tag{23}$$

$$\begin{aligned}
&= \left(\sum_{k=1}^n k \cdot (k-1) \cdot e^{-\lambda} \cdot \frac{\lambda^k}{k!} \right) + \lambda \\
&= e^{-\lambda} \cdot \left(\sum_{k=2}^n \frac{\lambda^k}{(k-2)!} \right) + \lambda \\
&= e^{-\lambda} \cdot \left(\sum_{k=2}^n \frac{\lambda^2 \cdot \lambda^{k-2}}{(k-2)!} \right) + \lambda \\
&= e^{-\lambda} \cdot \lambda^2 \cdot \underbrace{\left(\sum_{k=2}^n \frac{\lambda^{k-2}}{(k-2)!} \right)}_{e^{\lambda}} + \lambda \\
&= \lambda^2 + \lambda
\end{aligned} \tag{24}$$

(22) $k = 0 \Rightarrow \Sigma = 0$.

(23) Reescribimos $k^2 = (k \cdot ((k-1) + 1))$.

(24) $k = 1 \Rightarrow \Sigma = 0$

. Ahora termino calculando la **varianza de una Poisson**

$$\begin{aligned}
V(x) &= E(x^2) - E(x)^2 \\
&= \lambda^2 + \lambda - \lambda^2 \\
&= \lambda
\end{aligned}$$

En síntesis sea: $X \sim P(\lambda)$

- $P(k) = P(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!} \quad \forall k \in \mathbb{Z}^{\geq 0}$.
- $E(x) = V(x) = \lambda$.

3.6.2. Distribución de Poisson vista como una binomial

Se puede usar si: $N \geq 100, p \leq 0,01$ y $n.p \leq 20$

$$\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \approx e^{-\lambda} \cdot \frac{\lambda^k}{k!} : \lambda = n.p$$

Ejemplo

En una prueba de tarjetas de circuito la probabilidad de que un iodo falle es de 0,01. Suponiendo que la tarjeta tiene 200 de éstos, que funcionan independientemente uno de otro. ¿Cuál es la probabilidad de que fallen por lo menos 2 iodios en una tarjeta seleccionada al azar?

Sea X = “Número de iodios que fallan en una tarjeta”

$$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X = 0) - P(X = 1)$$

Si vemos a $X \sim B(200, 0,01)$

$$\begin{aligned} \text{Nos quedaría } P(X \geq 2) &= 1 - \binom{200}{0} \cdot p^0 \cdot (1-p)^{200} - \binom{200}{1} \cdot p^1 \cdot (1-p)^{199} \\ &= 1 - (0,99)^{200} - 200 \cdot (0,01) \cdot (0,99)^{199} \\ &= 0,5953 \end{aligned}$$

Si lo vemos como una Poisson de parámetro $\lambda = n.p = 200 * 0,01 = 2$

$$\begin{aligned} \text{Nos quedaría } P(X \geq 2) &= 1 - P(X < 2) \\ &= 1 - P(X = 0) - P(X = 1) \\ &\approx 1 - e^{-\lambda} \cdot \frac{\lambda^0}{0!} - e^{-\lambda} \cdot \frac{\lambda}{1!} \\ &\stackrel{\lambda=2}{=} 1 - 3 \cdot e^{-2} \\ &= 0,5940 \end{aligned}$$

3.7. Distribución hiper geométrica

Sea un experimento que cumpla las siguientes condiciones

- La población consta de N objetos o sujetos.
- Cada objeto o sujeto puede ser clasificado como éxito (E) o fracaso (F), y hay M éxitos en la población: $M \leq N$
- Se elige una muestra de tamaño n donde cada subconjunto tiene igual probabilidad de ocurrir.

Entonces la variable aleatoria X que cuenta el número de éxitos (E) de la muestra de tamaño n se llama hiper geométrica de parámetros n, M, N .

Notación: $X \sim H(n, M, N)$

3.7.1. Propiedades de la hiper geométrica

- $P(k) = P(X = k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}} : 0 \leq k < m \text{ y } 0 \leq n - k < N - M$
 $\therefore \max\{0, n - N + m\} \leq k \leq \min\{M, n\}$
- $E(x) = n \cdot \frac{M}{N}$
- $V(x) = n \cdot \underbrace{\frac{M}{N}}_p \cdot \underbrace{\left(1 - \frac{M}{N}\right)}_q \cdot \underbrace{\left(\frac{N-n}{N-1}\right)}_{\text{corrección}}$

Ejemplo

Cada uno de los 12 refrigeradores de cierto tipo ha sido devuelto a un distribuidor debido a la presencia de un ruido oscilante agudo cuando está funcionando. Suponiendo que 4 de esos 12 tienen compresores defectuosos y los otros 8 tienen problemas menos serios.

Definimos X = “número de refrigeradores que tienen el compresor defectuoso entre los primeros 6 examinados”

Si se examinan al azar, se pide calcular:

1. La probabilidad de que por lo menos uno tenga el compresor defectuoso.
2. La probabilidad de que haya entre 1 a 3 compresores rotos.

Comenzamos a analizar la variable aleatoria X y obtenemos que:

$$Im(X) = \{0, 1, 2, 3, 4\}$$

El tamaño de la muestra es $n = 6$ Sobre una población de $N = 12$
 La cantidad de éxitos es $M = 4$

Luego nos queda: $X \sim H(6, 4, 12)$

1. $P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) = 1 - \frac{\binom{8}{6}}{\binom{12}{6}}$
2. $P(1 \leq X \leq 3) = P(1) + P(2) + P(3) = 1 - (P(0) + P(4))$

$$= 1 - \frac{\binom{8}{6} + \binom{4}{4} \cdot \binom{8}{1}}{\binom{12}{6}}$$

Ejemplo

Un geólogo ha recolectado 10 especímenes de roca basáltica y 10 de rocas de granito. Se instruye a un ayudante de laboratorio para que seleccione al azar 15 de estos especímenes para analizarlos.

Sea X = “Número de rocas basálticas en la muestra”

Sea Y = “Número de rocas graníticas en la muestra”

Se pide:

1. Calcular la probabilidad de que todos los especímenes de uno de los dos tipos sean seleccionados.
2. Calcular la probabilidad de que X esté a menos de un desvío estándar de su valor medio.

El tamaño de la muestra es $n = 15$ Sobre una población de $N = 20$
 La cantidad de éxitos es $M = 10$

$$Im(X) = Im(Y) = \{5 \leq i \leq 10 : i \in \mathbb{N}\}$$

Tanto X como Y tienen distribución $X, Y \sim H(n, m, N) \sim H(15, 10, 20)$

$$1. P([X = 10] \cup [Y = 10]) = P([X = 10] \cup [X = 5]) = 2 \cdot \left(\frac{\binom{10}{10} \cdot \binom{10}{5}}{\binom{20}{15}} \right)$$

$$2. P(|X - \mu| < \sigma) = P(\mu - \sigma \leq X \leq \mu + \sigma)$$

$$\text{Sabemos que: } \mu = E(x) = n \cdot \frac{M}{N} = 15 \cdot \frac{10}{20} = 7,5$$

$$V(x) = \sigma^2 = n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right) \cdot \left(\frac{N-n}{N-1}\right) = 15 \cdot \frac{10}{20} \cdot \left(1 - \frac{10}{20}\right) \cdot \left(\frac{20-15}{20-1}\right)$$

$$\therefore \sigma = \sqrt{\frac{75}{76}}$$

$$6,51 = \mu - \sigma \leq X \leq \mu + \sigma = 8,49$$

$$\text{Luego } P(|X - \mu| < \sigma) = P(X = 7) + P(X = 8) = \frac{\binom{10}{7} \binom{10}{8} + \binom{10}{8} \binom{10}{7}}{\binom{20}{15}}$$

3.8. Distribución binomial negativa

Dado un experimento que cumpla:

- Los ensayos se realizan de forma independiente.
- Cada ensayo tiene dos resultados posibles (E) ó (F) y $P(E) = p$ para cada ensayo.
- Los ensayos se continúan hasta obtener r éxitos con $r \in \mathbb{N}$ fijo

Entonces la variable aleatoria X que cuenta el número de fracasos que preceden al r -ésimo éxito se llama binomial negativa de parámetros r y p .

Notación: $X \sim B^-(r, p) : Im(x) = \mathbb{Z}^{\geq 0}$

Definimos la función de probabilidad:

$$P(k) = P(X = k) = p^r \cdot (1 - p)^k \cdot \binom{k+r-1}{k}$$

Proposición

Sea $X \sim B^-(r, p)$ entonces:

- $E(x) = r \cdot \frac{(1-p)}{p}$
- $V(x) = r \cdot \frac{(1-p)}{p^2}$

Ejemplo

Volviendo al ejemplo de las baterías. Sea $Y =$ “número de baterías revisadas hasta obtener la primera que cumpla las condiciones de calidad”

$$\begin{aligned} Y &= X + 1 \\ E(Y) &= E(X) + 1 \\ \frac{1}{p} &= E(x) + 1 \\ \therefore E(X) &= \frac{1-p}{p} \end{aligned}$$

Ejemplo

Sea un estudio geológico en el cual se hacen perforaciones en la tierra hasta encontrar petróleo, y sea p la probabilidad de encontrarlo en cada perforación.

1. ¿Cuál es la probabilidad de que el primer descubrimiento ocurra en la tercera perforación?

Sea $X =$ “número de fracasos que preceden al primer éxito”

Definimos: $X \sim B^-(1, p)$ y calculamos $P(X=2)$

2. ¿Cuál es la probabilidad de que el tercer descubrimiento ocurra en la quinta perforación? Sea $Y =$ “número de fracasos que preceden al tercer éxito”

Definimos: $Y \sim B^-(3, p)$ y calculamos $P(Y=2)$

4. Variables aleatorias continuas

Definición: diremos que X es una variable aleatoria continua si:

$$P(X = x) = 0 \quad \forall x \in \mathbb{R}$$

4.1. Propiedades

Sea X una variable aleatoria entonces:

- $P(a \leq X \leq b) = F(b) - F(a)$
- $P(a \leq X \leq b) \stackrel{*}{=} P(a < X \leq b) \stackrel{*}{=} P(a \leq X < b)$
- $P(X > c) = 1 - P(X \leq c) = 1 - F(c) \stackrel{*}{=} P(x \geq c)$

★ Puede no valer si X es una variable aleatoria discreta.

Proposición

Sea un modelo probabilístico (S, \mathcal{A}, p)

- Sean $A_1 \subset A_2 \subset \dots$ Con $\{A_i\}_{i=1}^{\infty} : A_i \in \mathcal{A}$

$$\text{Entonces } P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} P(A_n)$$

Demostración: Defino: $\bigcup_{i=1}^{\infty} A_i \stackrel{*}{=} \bigcup_{i=1}^{\infty} (A_i - A_{i-1}) : A_0 = \emptyset$

★ Unión disjunta de eventos.

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i - A_{i-1}) \quad (25)$$

$$\begin{aligned} &= \lim_{n \rightarrow \infty} \sum_{i=1}^n P(A_i - A_{i-1}) \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n (P(A_i) - P(A_{i-1})) \end{aligned} \quad (26)$$

$$\begin{aligned} &= \lim_{n \rightarrow \infty} \sum_{i=1}^n P(A_i) - \lim_{n \rightarrow \infty} \sum_{i=1}^n P(A_{i-1}) \\ &= \lim_{n \rightarrow \infty} [P(A_1) + \dots + P(A_n) - (P(A_0) + \dots + P(A_{n+1}))] \\ &= \lim_{n \rightarrow \infty} [P(A_n) - P(A_0)] \\ &= \lim_{n \rightarrow \infty} P(A_n) \end{aligned} \quad (27)$$

(25) Disjuntos.

(26) Ya que $A_{i-1} \subset A_i \therefore P(A_i - A_{i-1}) = P(A_i) - P(A_{i-1})$.

(27) Ya que $P(A_0) = P(\emptyset) = 0$.

- Sean $B_1 \supset B_2 \supset \dots$ eventos de \mathcal{A} entonces: $P\left(\bigcap_{i=1}^{\infty} B_i\right) = \lim_{n \rightarrow \infty} P(B_n)$

$$\bigcap_{i=1}^{\infty} B_i = \overline{\left(\bigcap_{i=1}^{\infty} \overline{B_i}\right)} \quad (28)$$

$$= \overline{\left(\bigcup_{i=1}^{\infty} B_i\right)} \quad (29)$$

(28) Doble negación.

(29) De-Morgan.

$$\begin{aligned}
P\left(\bigcap_{i=1}^{\infty} B_i\right) &= \overline{P\left(\bigcup_{i=1}^{\infty} \overline{B_i}\right)} \\
&= 1 - P\left(\bigcup_{i=1}^{\infty} \overline{B_i}\right) \\
&= 1 - \lim_{n \rightarrow \infty} P(\overline{B_n}) \tag{30} \\
&= 1 - \lim_{n \rightarrow \infty} (1 - P(B_n)) \tag{31} \\
&= \lim_{n \rightarrow \infty} P(B_n)
\end{aligned}$$

(30) Vale por el item anterior.

(31) Definición de complemento.

4.2. Función de distribución acumulada

Sea (S, \mathcal{A}, p) un modelo probabilístico y X una V.A. entonces:

1. F es monótona creciente, id est: si $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$
2. $\lim_{x \rightarrow -\infty} F(x) = 0$ y el $\lim_{x \rightarrow \infty} F(x) = 1$
3. $\lim_{x \rightarrow a^+} F(x) = F(a)$
4. Sólo válida si es continua: $\lim_{x \rightarrow a^-} F(x) = \lim_{x \rightarrow a^+} F(x) = F(a)$

Demostración:

1. Si $x_1 < x_2 \Rightarrow [X \leq x_1] \subseteq [X \leq x_2]$

$$\therefore P(X \leq x_1) \leq P(X \leq x_2) = F(x_1) \leq F(x_2)$$

2. Para probar que el $\lim_{x \rightarrow \infty} F(x) = 1$ definamos $A_n = [X \leq n]$

$$\forall n \in \mathbb{N} \Rightarrow A_1 \subset A_2 \subset \dots \subset A_n \text{ y } \bigcup_{i=1}^{\infty} A_i = S \text{ espacio muestral}$$

$$1 = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} P(A_n) \stackrel{*}{=} \lim_{n \rightarrow \infty} P(X \leq n) = \lim_{n \rightarrow \infty} F(n)$$

$$\therefore \lim_{n \rightarrow \infty} F(n) = \lim_{x \rightarrow \infty} F(x) = 1$$

★ Por lo probado en la proposición anterior.

Ahora queremos probar que $\lim_{x \rightarrow -\infty} F(x) = 0$

Definimos $B_n = [X \leq -n] \forall n \in \mathbb{N}$

Ahora obtengo $B_1 \supset B_2 \supset B_3 \supset \dots$ Luego $\left(\bigcap_{n=1}^{\infty} B_n\right) = \emptyset$

$$0 = P\left(\underbrace{\bigcap_{n=1}^{\infty} B_n}_{P(\emptyset)}\right) \stackrel{*}{=} \lim_{n \rightarrow \infty} P(B_n) = \lim_{n \rightarrow \infty} F(-n) \therefore \lim_{x \rightarrow -\infty} F(x) = 0$$

★ Vale por lo probado en la proposición anterior.

3. Queremos demostrar que $\lim_{x \rightarrow a^+} F(x) = F(a)$

Definimos $B_n = [X \leq a + \frac{1}{n}] \forall n \in \mathbb{N}$

$$B_1 = [X \leq a + 1]$$

$$B_2 = [X \leq a + 0,5]$$

Luego $B_{i+1} \subset B_i \forall i : 1 \leq i \leq n$

Quiero ver que $\bigcap_{n=1}^{\infty} B_n = [X \leq a]$

(\subseteq) Sea $w \in \bigcap_{n=1}^{\infty} B_n \Rightarrow w \in B_n \forall n \in \mathbb{N}$

*asumimos $X(w) \not\leq a \Rightarrow X(w) \leq a + \frac{1}{n} \leq a \forall n \in \mathbb{N}$

$\therefore X(w) \leq a \Rightarrow w \in [X \leq a]$

(\supseteq) Sea $w \in [X \leq a] \Rightarrow X(w) \leq a \leq a + \frac{1}{n} \forall n \in \mathbb{N}$

$\Rightarrow w \in B_n \forall n \in \mathbb{N} \therefore w \in \bigcap_{n=1}^{\infty} B_n$

$\therefore \bigcap_{n=1}^{\infty} B_n = [X \leq a]$

Luego $F(a) = P\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n \rightarrow \infty} P(B_n) = \lim_{n \rightarrow \infty} F\left(a + \frac{1}{n}\right)$

$\therefore F(a) = \lim_{x \rightarrow a^+} F(x)$

4. Sea X una variable aleatoria continua queremos probar que:

$$\lim_{x \rightarrow a^-} F(x) = \lim_{x \rightarrow a^+} F(x) = F(a)$$

Definimos $A_n = [X \leq a - \frac{1}{n}] \quad \forall n \in \mathbb{N}$

Entonces $A_i \subset A_{i+1}$ y $\bigcup_{n=1}^{\infty} A_n \stackrel{*}{=} [X < a]$

★ Fácil de probar.

$$P(X < a) = P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{x \rightarrow \infty} P(A_n) = \lim_{x \rightarrow \infty} F(a - \frac{1}{n})$$

Como X es una variable aleatoria continua $P(X < a) = P(X \leq a)$

$$\text{Luego } P[X \leq a] = F(a) \Rightarrow \lim_{x \rightarrow a^-} F(x) = F(a) \stackrel{3}{=} \lim_{x \rightarrow a^+} F(x)$$

Corolario

- La función de distribución de probabilidad de una variable aleatoria continua es continua.
- La función de distribución de probabilidad acumulada de una variable aleatoria continua es continua.

Definición: se llama función de densidad de probabilidad a toda función

$$f : \mathbb{R} \rightarrow \mathbb{R}^{\geq 0} : \int_{-\infty}^{\infty} f(t) dt = 1$$

Propiedad: sea X una variable aleatoria continua entonces

$$f(t) = \begin{cases} F'(t) & \text{en donde exista} \\ 0 & \text{en caso contrario} \end{cases}$$

$$\text{y } F(x) = \int_{-\infty}^x f(t) dt$$

Definición percentil: sea X una variable aleatoria continua llamaremos percentil p : $(p(100\%))$ con $p \in (0, 1)$ al valor $\eta_p : F(\eta_p) = p$

Propiedad: sea X una variable aleatoria continua : $Y = a.X + b$ con $p \in (0, 1)$ y $a \neq 0$ entonces:

$$\eta_y(p) = \begin{cases} a.\eta_x(p) + b & \text{si } a > 0 \\ a.\eta_x(1 - p) + b & \text{si } a < 0 \end{cases}$$

Demostración

$$\begin{aligned} \text{Quiero hallar } \eta_y(p) : F_y(\eta_y(p)) &\stackrel{exijo}{=} p \stackrel{Def}{=} P(Y \leq \eta_y(p)) \\ &= P[a.X + b \leq \eta_y(p)] \end{aligned}$$

Si $a > 0$

$$\begin{aligned} [a.X + b \leq \eta_y(p)] &= \left[X \leq \frac{\eta_y(p) - b}{a} \right] \\ F_y(\eta_y(p)) = p &= F_x\left(\frac{\eta_y(p) - b}{a}\right) = F_x(\eta_x(p)) \end{aligned}$$

Luego $\eta_y(p) = a.\eta_x(p) + b$ si $a > 0$

Si $a < 0$

$$\begin{aligned} [a.X + b \leq \eta_y(p)] &= \left[X \geq \frac{\eta_y(p) - b}{a} \right] \\ F_y(\eta_y(p)) = p &= 1 - F_x\left(\frac{\eta_y(p) - b}{a}\right) = F_x(x) = 1 - p = F_x(1 - p) \\ \text{Luego } \eta_y(p) &= a.\eta_x(1 - p) + b \text{ si } a < 0 \end{aligned}$$

4.3. Distribución uniforme

Sea X una variable aleatoria será uniforme si su función de distribución acumulada está dada por:

$$F(x) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } x \in (a, b) \\ 1 & \text{si } x \geq b \end{cases} \quad \text{con } a < b$$

Como $F(x)$ no es diferenciable en $x = a$ y $x = b$ nos queda definida la función de distribución de probabilidad como:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in (a, b) \\ 0 & \text{en caso contrario} \end{cases}$$

Notación: $X \sim U(a, b)$

Propiedad: sea $X \sim U(a, b)$ queremos hallar $\eta_x(p) = p.(b - a) + a$

$$F(\eta_x(p)) = \frac{\eta_x(p) - a}{b - a} \Rightarrow \eta_x(p) = p.(b - a) + a$$

Con $p = 0,5 \Rightarrow \tilde{\mu} = \frac{1}{2}.(b - a) + a = \frac{b+a}{2}$ se define la mediana.

4.4. Esperanza y varianza de una variable aleatoria continua

Sea X una variable aleatoria continua con función de densidad de probabilidad f . Se define el valor esperado o valor medio o esperanza de X al valor medio.

Notación: $E(x) = \mu$

Proposición: sea X una variable aleatoria continua con función de densidad f y sea h una función entonces: $E(h(x)) = \int_{-\infty}^{\infty} h(x) \cdot f(x) dx$ siempre que exista.

Definición: llamaremos varianza de X a $E(x - \mu)^2$, siempre que exista donde $\mu = E(x)$ que denotaremos como $V(x) = \sigma^2$ y la desviación estándar de X como $\sqrt{V(x)}$.

Proposición: sea X una variable aleatoria con función de densidad f

- $V(x) = E(x^2) - E(x)^2$
- $E(a \cdot x + b) = a \cdot E(x) + b \quad \forall a, b \in \mathbb{R}$
- $V(a \cdot x + b) = a^2 \cdot V(x) \quad \forall a, b \in \mathbb{R}$

Demostración: Queremos ver que $V(x) = E(x^2) - E(x)^2$

$$\begin{aligned} V(x) &= E(\underbrace{(x - \mu)^2}_{h(x)}) = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x) dx \\ &= \int_{-\infty}^{\infty} x^2 \cdot f(x) dx + \underbrace{\mu^2 \cdot \int_{-\infty}^{\infty} f(x) dx}_{= 1} - 2 \cdot \underbrace{\mu \cdot \int_{-\infty}^{\infty} x \cdot f(x) dx}_{= \mu} \\ &= E(x^2) + \mu^2 - 2 \cdot \mu^2 \\ &= E(x^2) - \mu^2 \end{aligned}$$

4.4.1. Esperanza y varianza de una uniforme

Sea $X \sim U(0, 1) \Rightarrow E(x) = \frac{1}{2}$ y la $V(x) = \frac{1}{12}$

Demostración

- $E(x) = \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_0^1 x \cdot \underbrace{f(x)}_{=1} dx = \frac{x^2}{2} \Big|_0^1 = \frac{1}{2}$
- $E(x^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x) dx = \int_0^1 x^2 \cdot \underbrace{f(x)}_{=1} dx = \frac{x^3}{3} \Big|_0^1 = \frac{1}{3}$
- $V(x) = E(x^2) - E(x)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$

Proposición

$\forall a < b$ sea $Y = (b - a) \cdot X + a$ con $X \sim U(0, 1)$ entonces:

- $Y \sim U(a, b)$
- $E(Y) = \left(\frac{a+b}{2}\right)$
- $V(Y) = \frac{(b-a)^2}{12}$

Demostración:

$$F_y(y) = P(Y \leq y) = P\left(X \leq \frac{y-a}{b-a}\right) = F_x\left(\frac{y-a}{b-a}\right)$$

$$F'_y(y) = F'_x\left(\frac{y-a}{b-a}\right) \cdot \frac{1}{b-a} = f_x\left(\frac{y-a}{b-a}\right) \cdot \frac{1}{b-a}$$

$$f(y) = \begin{cases} \frac{1}{b-a} & \text{si } y \in (a, b) \\ 0 & \text{en caso contrario} \end{cases}$$

$$f(y) = I_{(a, b)}(y) \cdot \frac{1}{b-a} \Rightarrow Y \sim U(a, b)$$

Luego como $Y = (b - a) \cdot x + a$

$$E(x) = \frac{1}{2}$$

$$V(x) = \frac{1}{12}$$

Usando las propiedades de la esperanza y la varianza obtenemos que:

$$E(y) = (b - a) \cdot E(x) + a = \frac{b - a}{2} + a = \frac{b + a}{2}$$

$$V(y) = (b - a) \cdot V(x) = \frac{(b - a)^2}{12}$$

Corolario: sea $Y \sim U(-a, a)$ entonces $E(y) = 0$ y la $V(y) = \frac{a^2}{4}$

Ejemplo

El tiempo que tarda en realizar un viaje ida y vuelta un camión, que transporta concreto hacia una obra en construcción, tiene una distribución uniforme en el intervalo de 50 a 70 minutos.

1. ¿Cuál es la probabilidad que la duración del viaje sea mayor a 65 minutos dado que la duración ya pasó lo 55 minutos?

$$f(y) = \begin{cases} \frac{1}{b-a} & \text{si } y \in (a, b) \\ 0 & \text{en caso contrario} \end{cases}$$

$$f(x) = \begin{cases} \frac{1}{20} & \text{si } x \in (50, 70) \\ 0 & \text{en caso contrario} \end{cases}$$

$$\begin{aligned} P(X > 65 | X > 55) &= \frac{P((X > 65) \cap (X > 55))}{P(X > 55)} \\ &= \frac{P(X > 65)}{P(X > 55)} = \frac{1/4}{3/4} = \frac{1}{3} \end{aligned}$$

$$P(X > 65) = \int_{-\infty}^{\infty} f(x) dx = \int_{65}^{70} \frac{dx}{20} = \frac{x}{20} \Big|_{65}^{70} = \frac{1}{4}$$

$$P(X > 55) = \int_{-\infty}^{\infty} f(x) dx = \int_{55}^{70} \frac{dx}{20} = \frac{x}{20} \Big|_{55}^{70} = \frac{3}{4}$$

2. Dar el valor medio y la desviación estándar del tiempo de duración de viaje.

$$E(x) = \frac{a+b}{2} = \frac{70+50}{2} = 60$$

$$\sigma^2 = V(x) = \frac{(b-a)^2}{12} = \frac{20^2}{12} = 33,33$$

$$\sigma = \sqrt{V(x)} = \sqrt{33,33} = 5,7735$$

3. Suponga que los tiempos que tardan cada uno de los tres camiones son independientes uno de otro. ¿Cuál es la probabilidad de que exactamente uno de ellos tarde más de 55 minutos?

Defino una binomial: $X \sim B(n, p)$

Donde X es el número de camiones que tarda más de 55 minutos.

$$P(E) = p = \frac{3}{4} \therefore X \sim B\left(3, \frac{3}{4}\right)$$

Me pide calcular la $P(X = 1) = \binom{3}{1} \cdot \frac{3}{4} \cdot \left(\frac{1}{4}\right)^{(3-1)} = 3 \cdot \frac{3}{4} \cdot \left(\frac{1}{4}\right)^2 = 0,1406$

4.5. Distribución normal

Definición: diremos que una variable aleatoria X tiene distribución normal de parámetros $\mu \in \mathbb{R}$ y $\sigma^2 > 0$ si su función de densidad de probabilidad está dada por:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \forall x \in \mathbb{R}$$

Es la variable aleatoria continua más importante dentro de la probabilidad y estadística.

$$\text{Notación : } X \sim N(\mu, \sigma^2)$$

4.5.1. Propiedades de la función de densidad

- f es simétrica en torno de μ , id est: $f(\mu - t) = f(\mu + t) \quad \forall t \in \mathbb{R}$.
- f tiene un punto máximo en $x = \mu$.
- f tiene puntos de inflexión en $\mu + \sigma$ y $\mu - \sigma$.
- $\lim_{x \rightarrow -\infty} f(x) = \lim_{x \rightarrow \infty} f(x) = 0$

Proposición

Sea $X \sim N(\mu, \sigma^2)$ entonces:

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

Prueba:

$$\int_{-\infty}^{\infty} f(x) dx = 1 \Leftrightarrow \left[\int_{-\infty}^{\infty} f(x) dx \right]^2 = 1$$

hacemos un cambio de variable llamando a: $\frac{x-\mu}{\sigma} = t$

como $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left(\frac{x-\mu}{\sigma}\right)^2 \frac{1}{2}}$ obtengo $f(t) = \frac{\sigma}{\sqrt{2\pi\sigma^2}} e^{-\frac{t^2}{2}}$

$$\begin{aligned}
\text{Partimos de } \int_{-\infty}^{\infty} \frac{\sigma}{\sqrt{2\pi\sigma^2}} e^{\frac{-t^2}{2}} dt &= \left[\int_{-\infty}^{\infty} \frac{\sigma}{\sqrt{2\pi\sigma^2}} e^{\frac{-t^2}{2}} dt \right]^2 \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\frac{-t^2}{2}} dt \cdot \int_{-\infty}^{\infty} e^{\frac{-y^2}{2}} dy \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{\frac{-t^2-y^2}{2}} dt dy \\
&= \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} e^{\frac{-r^2}{2}} \cdot r dr d\theta \quad (32)
\end{aligned}$$

$$= \frac{1}{2\pi} \int_0^{2\pi} d\theta \quad (33)$$

$$= 1 \quad (34)$$

(32) Hacemos un cambio a coordenadas polares llamando a:

$$t = r \cdot \cos(\theta)$$

$$y = r \cdot \sin(\theta) : 0 \leq \theta \leq 2\pi \Rightarrow \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} e^{\frac{-r^2}{2}} \cdot r dr d\theta$$

$$(33) \int_0^{\infty} e^{\frac{-r^2}{2}} \cdot r dr = \left(-e^{\frac{-r^2}{2}} \right) \Big|_0^{+\infty} = \left(\lim_{r \rightarrow \infty} -e^{\frac{-r^2}{2}} \right) + 1 = 1$$

$$(34) \frac{1}{2\pi} \int_0^{2\pi} d\theta = \frac{1}{2\pi} (\theta \Big|_0^{2\pi}) = 2\pi \cdot \frac{1}{2\pi} = 1$$

2. $\boxed{E(x) = \mu}$

$$E(x) = \int_{-\infty}^{\infty} ((t - \mu) + \mu) \cdot f(t) dt \quad (35)$$

$$\begin{aligned} &= \int_{-\infty}^{\infty} (t - \mu) \cdot f(t) dt + \int_{-\infty}^{\infty} \mu \cdot f(t) dt \\ &= \int_{-\infty}^{\infty} (t - \mu) \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\left(\frac{t-\mu}{\sigma}\right)^2 \frac{1}{2}} dt + \mu \cdot \int_{-\infty}^{\infty} f(t) dt \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{t - \mu}{\sigma}\right) \cdot e^{-\left(\frac{t-\mu}{\sigma}\right)^2 \frac{1}{2}} dt + \mu \end{aligned} \quad (36)$$

$$= \mu \quad (37)$$

(35) Sumo y resto μ

(36) Ya que $\int_{-\infty}^{\infty} f(t) dt = 1$

(37) $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{t-\mu}{\sigma}\right) e^{-\left(\frac{t-\mu}{\sigma}\right)^2 \frac{1}{2}} dt = 0$

3. $V(x) = \sigma^2$

$$\begin{aligned}
 V(x) &= E(x - \mu)^2 \\
 &= \int_{-\infty}^{\infty} \left(\frac{x - \mu}{\sigma} \right)^2 \cdot \frac{\sigma^2}{\sqrt{2\pi}} \cdot e^{-\left(\frac{x-\mu}{\sigma}\right)^2/2} dx \\
 &= \int_{-\infty}^{\infty} y^2 \frac{\sigma^2}{\sqrt{2\pi}} \cdot e^{-\frac{y^2}{2}} dy \tag{38}
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 \cdot e^{-\frac{y^2}{2}} dy \\
 &= \frac{\sigma^2}{\sqrt{2\pi}} \left(-y \cdot e^{-\frac{y^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy \right) \\
 &= \sigma^2 \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy \right] \tag{39} \\
 &= \sigma^2
 \end{aligned}$$

(38) Tomo $y = \frac{x-\mu}{\sigma} \Rightarrow \frac{dy}{dx} = \frac{1}{\sigma}$

(39) $-y \cdot e^{-\frac{y^2}{2}} \Big|_{-\infty}^{\infty} = - \lim_{y \rightarrow -\infty} \frac{y}{e^{\frac{y^2}{2}}} + \lim_{y \rightarrow \infty} \frac{y}{e^{\frac{y^2}{2}}} = 0$

Dentro de la familia de normales la más importante es cuando $\mu = 0$ y $\sigma = 1$ tal variable se denomina **variable normal estándar**.⁵

⁵Notación: $Z \sim N(0, 1)$

4.5.2. Estandarización

Sea $X \sim N(\mu, \sigma^2) \Rightarrow \left(\frac{X-\mu}{\sigma}\right) \sim N(0, 1)$

Demostración: Sea $Y = \left(\frac{x-\mu}{\sigma}\right)$

$$\begin{aligned} F_y(y) &= P[Y \leq y] \\ &= P\left[\frac{x-\mu}{\sigma} \leq y\right] \\ &= P[x \leq \sigma \cdot y + \mu] \\ &= F_x(\sigma \cdot y + \mu) \quad \forall y \in \mathbb{R} \\ &= \sigma \cdot f_x(\sigma \cdot y + \mu) \\ &= \frac{\sigma}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\sigma \cdot y + \mu)^2}{2\sigma^2}} \\ &= \frac{e^{-\frac{\sigma^2 \cdot y^2}{2\sigma^2}}}{\sqrt{2\pi}} \\ &= \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \quad \forall y \in \mathbb{R} \quad \therefore Y \sim N(0, 1) \end{aligned}$$

Luego $P(a \leq X \leq b) \stackrel{Estand.}{=} P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$
donde Φ es la función de distribución acumulada de Z .

Ejemplo

Sea $Z \sim N(0, 1)$ se pide:

1. $P(Z \leq 2) = \Phi(2) = 0,9772$
2. $P(Z \geq 1,96) = 1 - \Phi(1,96) = 1 - 0,9750 = 0,025$
3. $P(0 \leq Z \leq 1,73) = \Phi(1,73) - \Phi(0) = 0,9582 - 0,5 = 0,4582$
4. $P(-1,5 \leq Z \leq -1) = \Phi(1,5) - \Phi(1) = 0,9332 - 0,8413 = 0,0919$
5. $P(-1,5 \leq Z \leq 2,82) = \Phi(2,82) - \Phi(-1,5) = \Phi(2,82) - (1 - \Phi(1,5))$
 $= 0,9976 - 1 + 0,9332$
 $= 0,9308$

Propiedad

Sea $X \sim N(\mu, \sigma^2)$ y $p \in (0,1) \Rightarrow \eta_x(p) = \sigma \cdot \eta_z(p) + \mu$

Demostración

$$p = F_x(\eta_x(p)) = P(X = \eta_x(p)) \stackrel{Estand.}{=} P\left(Z \leq \frac{\eta_x(p) - \mu}{\sigma}\right) = \Phi\left(\frac{\eta_x(p) - \mu}{\sigma}\right)$$

$$\text{Luego } \eta_z(p) = \frac{\eta_x(p) - \mu}{\sigma} \therefore \eta_x(p) = \sigma \cdot \eta_z(p) + \mu .$$

Ejemplo

Los alambres que se utilizan en cierta computadora deben tener una resistencia $\Omega \in (0,12 ; 0,14)$. Para cierta compañía la resistencia de los alambres son producidos con una distribución normal con una media de 0,13 y una desviación estándar $\sigma = 0,005$.

Se pide:

1. ¿Cuál es la probabilidad que un alambre producido en dicha compañía cumpla con las especificaciones?

$$\begin{aligned} P(0,12 \leq X \leq 0,14) &= P\left(\frac{0,12 - 0,13}{0,005} \leq Z \leq \frac{0,14 - 0,13}{0,005}\right) \quad (40) \\ &= P(-2 \leq Z \leq 2) \\ &= \Phi(2) - (1 - \Phi(2)) \\ &= 2 \cdot \Phi(2) - 1 \\ &= 0,9544 \end{aligned}$$

(40) Estandarizo

2. Si se seleccionan 4 alambres para un sistema al azar ¿Cuál es la probabilidad que los 4 cumplan con las especificaciones?

$$(0,9544)^4 = 0,8297$$

3. ¿Cuál es la probabilidad de que exactamente una cumpla con las especificaciones?

Defino una binomial $Y : Y \sim B(4; 0,9544)$

entonces me pide calcular $P[Y = 1] = \binom{4}{1} \cdot (1 - p)^3 \cdot p = 0,13$

4. Para qué valor c se cumple que la resistencia de un alambre se desvíe a lo sumo c unidades con un porcentaje del 95 %.

$$P(|x - \mu| \leq c) \stackrel{exijo}{=} 0,95$$

$$\text{Estandarizo y obtengo: } P\left(\frac{X-\mu}{\sigma} \leq \frac{c}{\sigma}\right) \stackrel{exijo}{=} 0,95$$

$$P(Z \leq z) : z = \frac{c}{\sigma} = 1,96 \quad \therefore c = 1,96 * \sigma = 0,0098$$

4.6. Distribución Gamma

Sea $f : \mathbb{R}^{\geq 0} \rightarrow \mathbb{R}^{\geq 0}$ llamaremos función gamma a la definida por:

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx \quad \forall \alpha > 0$$

1. $\Gamma(1) = 1$
2. $\Gamma(\alpha + 1) = \alpha \cdot \Gamma(\alpha)$
3. $\Gamma(n + 1) = n! \quad \forall n \in \mathbb{N}$

Demostración:

$$\begin{aligned} \Gamma(1) &= \int_0^{+\infty} x^{1-1} \cdot e^{-x} dx \\ &= \int_0^{+\infty} e^{-x} dx \\ &= -e^{-x} \Big|_0^{\infty} \\ &= \lim_{x \rightarrow \infty} \frac{1}{e^x} + 1 \\ &= 1 \end{aligned}$$

Ahora quiero ver que $\Gamma(\alpha + 1) = \alpha.\Gamma(\alpha)$

$$\begin{aligned}
\Gamma(\alpha) &= \int_0^{+\infty} x^{\alpha-1} \cdot e^{-x} dx \\
&= \int_0^{+\infty} x \cdot x^{\alpha-2} \cdot e^{-x} dx \\
&= \int_0^{+\infty} \underbrace{x}_{U'} \cdot \underbrace{x^{\alpha-2} e^{-x}}_V dx = \underbrace{\left(\frac{x^2}{2} x^{\alpha-2} e^{-x} \right) \Big|_0^{\infty}}_{=0} \\
&\quad - \frac{1}{2} \left[(\alpha - 2) \cdot \int_0^{+\infty} x^2 x^{\alpha-3} e^{-x} dx - \int_0^{+\infty} x^2 x^{\alpha-2} e^{-x} dx \right] \\
&= - \left(\frac{\alpha - 2}{2} \right) \left(\underbrace{\int_0^{+\infty} x^2 x^{\alpha-3} e^{-x} dx}_{\Gamma(\alpha)} - \underbrace{\int_0^{+\infty} x^2 x^{\alpha-2} e^{-x} dx}_{\Gamma(\alpha+1)} \right) \\
&= - \left(\frac{\alpha - 2}{2} \right) \cdot \Gamma(\alpha) + \frac{\Gamma(\alpha + 1)}{2} \\
\frac{\Gamma(\alpha + 1)}{2} &= \Gamma(\alpha) \cdot \left[1 + \left(\frac{\alpha - 2}{2} \right) \right] \\
\frac{\Gamma(\alpha + 1)}{2} &= \frac{\alpha}{2} \cdot \Gamma(\alpha)
\end{aligned} \tag{41}$$

$$\therefore \Gamma(\alpha + 1) = \alpha.\Gamma(\alpha) \quad \blacksquare$$

$$(41) \text{ Entonces } U = \frac{x^2}{2} \text{ y } V' = (\alpha - 2) \cdot x^{\alpha-3} e^{-x} - e^{-x} x^{\alpha-2}$$

Por último quiero ver que $\Gamma(n + 1) = n! \forall n \in \mathbb{N}$

Prueba por inducción en n:

Si $n = 0 \Rightarrow \Gamma(0 + 1) = \Gamma(1) = 1 = 0!$ Vale

Tomo como hipótesis inductiva que vale para n, id est: $\Gamma(n + 1) = n!$

Veamos que $\Gamma((n + 1) + 1) = (n + 1)!$

$$\Gamma((n + 1) + 1) \stackrel{(2)}{=} (n + 1) \cdot \Gamma(n + 1) \stackrel{H.I.}{=} (n + 1) \cdot n! = (n + 1)!$$

Definición: diremos que una variable aleatoria X tiene distribución Gamma de parámetros α y β con $(\alpha, \beta > 0)$ si tiene función de densidad definida por:

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha) \cdot \beta^\alpha} \cdot x^{\alpha-1} \cdot e^{-\frac{x}{\beta}} & \text{si } x > 0 \\ 0 & \text{en caso contrario} \end{cases}$$

Notación: $X \sim \Gamma(\alpha, \beta)$

Ahora probaremos que la integral de la función de densidad da 1

Id est: $\int_{-\infty}^{\infty} f(x) dx \stackrel{?}{=} 1$

Como la función de densidad sólo está definida si $x > 0$

$$\begin{aligned} \frac{1}{\Gamma(\alpha) \beta^\alpha} \int_0^{+\infty} x^{\alpha-1} e^{-\frac{x}{\beta}} dx &= \frac{1}{\Gamma(\alpha) \beta^{\alpha-1} \cdot \beta} \int_0^{+\infty} x^{\alpha-1} e^{-\frac{x}{\beta}} dx \\ &= \frac{1}{\Gamma(\alpha) \cdot \beta} \int_0^{+\infty} \left(\frac{x}{\beta}\right)^{\alpha-1} \cdot e^{-\frac{x}{\beta}} dx \\ &= \frac{1}{\Gamma(\alpha) \cdot \beta} \underbrace{\int_0^{+\infty} t^{\alpha-1} e^{-t} dt}_{\Gamma(\alpha)} \cdot \beta \quad (42) \\ &= \frac{1}{\Gamma(\alpha)} \cdot \Gamma(\alpha) \\ &= 1 \end{aligned}$$

(42) Hago un cambio de variable llamando a: $t = \frac{x}{\beta}$

Luego queremos Hallar $E(x)$ y la $V(x)$

Si $X \sim \Gamma(\alpha, \beta)$ y sea $n \in \mathbb{N}$

$$\begin{aligned}
 E(x^n) &= \int_{-\infty}^{+\infty} x^n f(x) dx \\
 &= \int_0^{\infty} x^n \cdot \frac{x^{\alpha-1}}{\Gamma(\alpha) \cdot \beta^\alpha} \cdot e^{-\frac{x}{\beta}} dx \\
 &= \frac{\beta^{n-1}}{\Gamma(\alpha)} \int_0^{\infty} \left(\frac{x}{\beta}\right)^{(n+\alpha)-1} \cdot e^{-\frac{x}{\beta}} dx \\
 &= \frac{\beta^{n-1} \cdot \beta}{\Gamma(\alpha)} \int_0^{\infty} t^{(n+\alpha)-1} \cdot e^{-t} dt \\
 &= \frac{\beta^n \cdot \Gamma(n+\alpha)}{\Gamma(\alpha)}
 \end{aligned} \tag{43}$$

(43) Hago un cambio de variables llamando a: $t = \frac{x}{\beta}$

$$\begin{aligned}
 E(x) &= \frac{\beta \cdot \Gamma(\alpha+1)}{\Gamma(\alpha)} \\
 &= \beta \cdot \alpha
 \end{aligned}$$

$$\begin{aligned}
 V(x) &= E(x^2) - (E(x))^2 \\
 &= \frac{\beta^2 \Gamma(\alpha+2)}{\Gamma(\alpha)} - \beta^2 \alpha^2 \\
 &= \beta^2 \cdot \left(\frac{\alpha \cdot (\alpha+1) \Gamma(\alpha)}{\Gamma(\alpha)} \right) - \beta^2 \alpha^2 \\
 &= \beta^2 \alpha^2 + \beta^2 \alpha - \beta^2 \alpha^2 = \beta^2 \alpha \\
 &= \beta^2 \alpha
 \end{aligned}$$

4.7. Distribución exponencial

Definición: diremos que X tiene distribución exponencial de parámetro λ (con $\lambda > 0$) si su función de densidad está dada por:

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{en caso contrario} \end{cases}$$

4.7.1. Distribución exponencial vista como una Gamma

$$\text{Si } X \sim \Gamma(1, \beta) \Rightarrow f(x) = \begin{cases} \frac{1}{\beta} \cdot e^{-\frac{x}{\beta}} & \text{si } x > 0 \\ 0 & \text{en caso contrario} \end{cases}$$

$$\text{Luego llamando } \lambda = \frac{1}{\beta} \Rightarrow f(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{en caso contrario} \end{cases}$$

4.7.2. Propiedades de la distribución exponencial

Sea $X \sim \exp(\lambda) = \Gamma\left(1, \frac{1}{\lambda}\right)$ vale que:

1. $E(x) = \frac{1}{\lambda}$ y la $V(x) = \frac{1}{\lambda^2}$

2. $F(x) = P[X \leq x] = \int_{-\infty}^{+\infty} f(t) dt$

$$\underline{\text{Si } x \leq 0} \Rightarrow F(x) = 0$$

$$\begin{aligned} \underline{\text{Si } x > 0} \Rightarrow F(x) &= \int_0^x \lambda \cdot e^{-\lambda \cdot t} dt \\ &= -e^{-\lambda \cdot t} \Big|_0^x \\ &= -e^{-\lambda \cdot x} + 1 \end{aligned}$$

$$\text{Luego } F(x) = \begin{cases} 1 - e^{-\lambda \cdot x} & \text{si } x > 0 \\ 0 & \text{en caso contrario} \end{cases}$$

3. Propiedad de falta de memoria

$$P(x \geq t + t_0 | x \geq t_0) = P(x \geq t)$$

$$\begin{aligned} P(x \geq t + t_0 | x \geq t_0) &= \frac{P((x \geq t + t_0) \cap (x \geq t_0))}{P(x \geq t_0)} \\ &= \frac{P(x \geq t + t_0)}{P(x \geq t_0)} \\ &= \frac{1 - F(t + t_0)}{1 - F(t_0)} \\ &= \frac{1 - (1 - e^{-\lambda(t+t_0)})}{1 - (1 - e^{-\lambda.t_0})} \\ &= \frac{e^{-\lambda(t+t_0)}}{e^{-\lambda.t_0}} \\ &= e^{-\lambda.t} \\ &= 1 - (1 - e^{-\lambda.t}) \\ &= P(x \geq t) \end{aligned}$$

4.8. Distribución χ^2

Definición: diremos que una variable aleatoria X tiene distribución χ^2 o chi cuadrado de parámetro ν , ($\nu > 0$) si su función de densidad es la de $\Gamma\left(\frac{\nu}{2}, 2\right)$

$$\text{Notación: } X \sim \chi_\nu^2 = \Gamma\left(\frac{\nu}{2}, 2\right)$$

4.8.1. Propiedades de la distribución χ^2

Sea $X \sim \chi_\nu^2 = \Gamma\left(\frac{\nu}{2}, 2\right)$ vale que:

1. $E(x) = \frac{\nu}{2}, 2 = \nu$ y la $V(x) = 2\nu$
2. Sea $Z \sim N(0, 1) \Rightarrow Z^2 \sim \chi_1^2 = \Gamma\left(\frac{1}{2}, 2\right)$

Demostración:

Si $w \leq 0 \Rightarrow F_w(w) = P(Z^2 \leq w) = P(\emptyset) = 0$

Si $w > 0$ $F_w(w) = P(Z^2 \leq w) = P(|z| \leq \sqrt{w}) = \Phi(\sqrt{w}) - \Phi(-\sqrt{w})$

Donde Φ es la función de distribución acumulada de la normal estándar.

$$\begin{aligned} F'_w(w) &= f_w(w) = f_z(\sqrt{w})\left(\frac{1}{2\sqrt{w}}\right) - f_z(-\sqrt{w})\left(-\frac{1}{2\sqrt{w}}\right) \\ &= \frac{1}{2\sqrt{w}} \left[\frac{1}{\sqrt{2\pi}} [e^{-\frac{w}{2}} + e^{-\frac{w}{2}}] \right] \\ &= \frac{e^{-\frac{w}{2}}}{\sqrt{w}\sqrt{2\pi}} = \frac{1}{w^{1/2}\sqrt{2\pi}} e^{-\frac{w}{2}} \\ &= \frac{w^{-1/2}}{\sqrt{2\pi}} e^{-\frac{w}{2}} = \frac{w^{-1/2}}{2^{1/2}\sqrt{\pi}} e^{-\frac{w}{2}} \end{aligned}$$

Luego si probamos $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \Rightarrow z^2 \approx \chi_1^2$

★ Recordar que si $Z \sim N(0, 1) \Rightarrow f_z(w) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$

Mis candidatos van a ser $\beta = 2$ y $\alpha = \frac{1}{2}$

$$1 = \int_0^{\infty} f_w(w) dw \quad (44)$$

$$= \frac{2^{-\frac{1}{2}}}{2^{\frac{1}{2}} \cdot \sqrt{\pi}} \int_0^{\infty} \left(\frac{w}{2}\right)^{-1/2} e^{-w/2} dw \quad (45)$$

$$= \frac{1}{2\sqrt{\pi}} \int_0^{\infty} 2 \cdot t^{-1/2} e^{-t} dt \quad (46)$$

$$= \frac{1}{\sqrt{\pi}} \underbrace{\int_0^{\infty} t^{-1/2} e^{-t} dt}_{\Gamma(\frac{1}{2})}$$

$$\therefore \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

(44) Por ser $f_w(w)$ una función de densidad.

(45) Multiplico y divido por $2^{-\frac{1}{2}}$

(46) Hago un cambio de variables llamando a: $t = \frac{w}{2}$

5. Distribución de probabilidad conjunta

Sean X e Y dos variables aleatorias sobre un mismo espacio muestral S . Definimos la función de probabilidad conjunta de (x, y) como la función:

$$P : \mathbb{R}^2 \rightarrow [0, 1]$$

Dada por: $P(x, y) = P([X = x] \cap [Y = y]) \quad \forall x, y \in \mathbb{R}$

A partir de la función de probabilidad conjunta se puede hallar la función de probabilidad de X e Y que se las conoce con el nombre de marginales.

$$\begin{aligned} P_x(x) &= P(X = x) = P([X = x] \cap [\cup_y [Y = y]]) \\ &= P(\cup_y [X = x] \cap [Y = y]) \\ &= \sum_Y \underbrace{P([X = x] \cap [Y = y])}_{P(x, y)} \\ \therefore P_x(x) &= \sum_Y P(x, y) \quad \forall x \in \mathbb{R} \end{aligned}$$

De manera similar resulta que: $P_y(y) = P(Y = y) = \sum_X P(x, y) \quad \forall y \in \mathbb{R}$

5.1. Funciones marginales

Sean X e Y variables aleatorias discretas se cumple que:

La marginal de x es: $P_x(x) = \sum_Y P(x, y) \quad \forall x \in \mathbb{R}$

La marginal de y es: $P_y(y) = \sum_X P(x, y) \quad \forall y \in \mathbb{R}$

Ejemplo

Se asignan aleatoriamente 2 contratos de construcciones para 3 empresas posibles: A,B,C.

Sea X = “Número de contratos asignados a la empresa A”

Sea Y = “Número de contratos asignados a la empresa B”

Los valores posibles de X e Y son: 0, 1, 2.

$S = \{(A,A);(A,B);(A,C);(B,A);(B,B);(B,C);(C,A);(C,B);(C,C)\}$

Si todos los pares son igualmente posible entonces $P(\{w\}) = \frac{1}{9} \quad \forall w \in S$

X \ Y	Y		
	0	1	2
0	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{9}$
1	$\frac{2}{9}$	$\frac{2}{9}$	0
2	$\frac{1}{9}$	0	0

Luego la marginal de x : $P_x(x) = \sum_{y=0}^2 P(x, y) = \frac{4}{9}$

$$\text{Luego } P_y(y) = P_x(x) = \begin{cases} \frac{1}{9} & \text{si } x = 2 \\ \frac{4}{9} & \text{si } x = 1 \text{ ó } x = 0 \\ 0 & \text{en caso contrario} \end{cases} \quad (\text{matriz simétrica}).$$

Como tienen la misma distribución entonces tienen la misma esperanza y varianza.

$$E(x) = E(y) = \sum_{y=0}^2 P[Y = y] = 0 * \frac{4}{9} + 1 * \frac{4}{9} + 2 * \frac{1}{9} = \frac{2}{3}$$

$$E(x^2) = E(y^2) = \sum_{y=0}^2 P[Y = y] = \sum_{y=1}^2 P[Y = y] = 1^2 * \frac{4}{9} + 2^2 * \frac{1}{9} = \frac{8}{9}$$

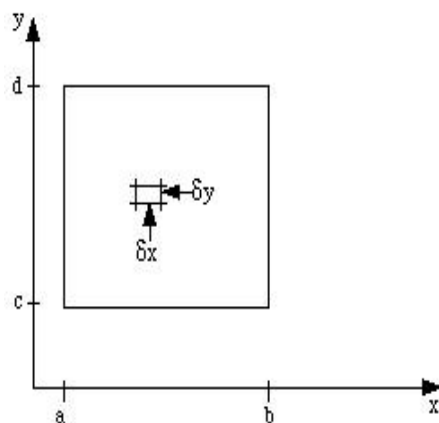
$$V(x) = V(y) = E(y^2) - E(y)^2 = \frac{8}{9} - \frac{4}{9} = \frac{4}{9} \quad \blacksquare$$

5.2. Función de densidad conjunta

Definición: sean X e Y variables aleatorias continuas sobre el mismo espacio muestral S . Llamaremos función de densidad conjunta de (x, y) a la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}^{\geq 0}$

$$\text{Tal que: } P((x, y) \in A) = \int_{(x,y)} \int_{(x,y)} f(x, y) \, dx dy$$

$$\text{Luego } P((x, y) \in [a, b] * [c, d]) = \int_c^d \int_a^b f(x, y) \, dx dy$$



$$\text{Si } A = \mathbb{R}^2 \Rightarrow P((x, y) \in \mathbb{R}^2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx dy = 1$$

$$\text{La marginal de X es: } f_x(x) = \int_{-\infty}^{\infty} f(x, y) \, dy \quad \forall x \in \mathbb{R}$$

$$\text{La marginal de Y es: } f_y(y) = \int_{-\infty}^{\infty} f(x, y) \, dx \quad \forall y \in \mathbb{R}$$

Ejemplo

Sea X = “Proporción de tiempo que está ocupada la casilla de autos”

Sea Y = “Proporción de tiempo que está ocupada la casilla de colectivos”

$$\text{Donde: } f(x, y) = \begin{cases} k(x + y^2) & \text{si } x, y \in [0, 1] \\ 0 & \text{en caso contrario} \end{cases}$$

- Determinar el valor positivo de k:

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

Reemplazo por la función, los límites de integración donde está definida y obtengo:

$$1 = \int_0^1 \int_0^1 k \cdot (x + y^2) \, dx dy = k \cdot \left(\int_0^1 \frac{x^2}{2} \Big|_0^1 + \int_0^1 y^2 x \Big|_0^1 \right) dy$$

$$1 = k \cdot \left(\int_0^1 \frac{1}{2} + y^2 \, dy \right) \Rightarrow k = \frac{6}{5}$$

$$\text{Luego } f(x, y) = \begin{cases} \frac{6}{5} * (x + y^2) & \text{si } x, y \in [0, 1] \\ 0 & \text{en caso contrario} \end{cases}$$

- Obtener la marginal de X:

$$\text{Si } x \notin [0, 1] \Rightarrow f_x(x) = \int_{-\infty}^{\infty} \underbrace{f(x, y)}_{=0} \, dy = 0$$

$$\begin{aligned} \text{Si } x \in [0, 1] \Rightarrow f_x(x) &= \int_{-\infty}^{\infty} f(x, y) \, dy \\ &= \frac{6}{5} \int_0^1 x + y^2 \, dy \\ &= \vdots \\ &= \frac{2}{5} \cdot (3x + 1) \end{aligned}$$

$$\text{Luego } f_x(x) = \begin{cases} \frac{2}{5} \cdot (3x + 1) & \text{si } x \in [0, 1] \\ 0 & \text{en caso contrario} \end{cases}$$

- Obtener la marginal de Y:

Se obtiene de la misma manera que la marginal de x pero integrando sobre x:

$$f_y(y) = \begin{cases} \frac{3}{5} \cdot (1 + y^2) & \text{si } y \in [0, 1] \\ 0 & \text{en caso contrario} \end{cases}$$

- Calcular la $E(x)$ y la $V(x)$

$$E(x) = \int_{-\infty}^{\infty} x \cdot f_x(x) dx = \frac{2}{5} \int_0^1 x \cdot (1 + 3x) dx = \frac{2}{5} \cdot \left[\frac{1}{2} + 3 \cdot \frac{1}{3} \right] = \frac{3}{5} = 0,6$$

$$E(x^2) = \int_{-\infty}^{\infty} x^2 \cdot f_x(x) dx = \frac{2}{5} \int_0^1 x^2 \cdot (1 + 3x) dx = \frac{2}{5} \cdot \left[\frac{1}{3} + 3 \cdot \frac{1}{4} \right] = \frac{13}{30} = 0,43$$

$$V(x) = E(x^2) - E(x)^2 = \frac{13}{30} - \frac{9}{25} = \frac{11}{150} = 0,073$$

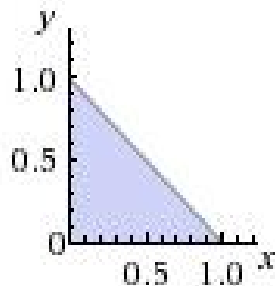
- Sea $A = \{(x, y) : x, y \leq \frac{1}{4}\}$ se pide calcular $P((x, y) \in A)$

$$P((x, y) \in A) = \int_0^{\frac{1}{4}} \int_0^{\frac{1}{4}} f(x, y) dx dy = \dots = \frac{7}{640} = 0,010$$

Ejemplo

$$\text{Sea } f(x, y) = \begin{cases} k \cdot x \cdot y & \text{si } x, y \in [0, 1] \\ 0 & \text{en caso contrario} \end{cases} : x + y \leq 1 \Rightarrow y \leq 1 - x$$

- Determinar el valor positivo de k



$$\begin{aligned}
1 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx dy \\
&= \int_0^1 \int_0^{1-y} k \cdot x \cdot y \, dx dy \\
&= k \cdot \int_0^1 y \left[\int_0^{1-y} x \, dx \right] dy \\
&= k \cdot \int_0^1 y \cdot \frac{(1-y)^2}{2} \, dy
\end{aligned}$$

$$\Leftrightarrow \frac{k}{2} \left[\frac{1}{2} + \frac{1}{4} - 2 \cdot \frac{1}{3} \right] = 1 \Leftrightarrow \frac{k}{24} = 1 \Leftrightarrow \boxed{k = 24}$$

- Obtener la marginal de X

$$\text{Si } x \notin [0, 1] \Rightarrow f_x(x) = \int_{-\infty}^{\infty} \underbrace{f(x, y)}_{=0} \, dy = 0$$

$$\text{Si } x \in [0, 1] \Rightarrow f_x(x) = \int_{-\infty}^{\infty} f(x, y) \, dy$$

$$\int_0^{1-x} f(x, y) \, dy = 24 \cdot x \cdot \int_0^{1-x} y \, dy = 24 \cdot x \cdot \frac{(1-x)^2}{2}$$

$$\text{Luego } f_x(x) = \begin{cases} 12 \cdot x \cdot (1-x)^2 & \text{si } x \in [0, 1] \\ 0 & \text{en caso contrario} \end{cases}$$

5.3. Interdependencia entre variables aleatorias

Si A y B son eventos independientes en el mismo espacio muestral S, defino que A y B son independientes si: $P(A \cap B) = P(A) \cdot P(B)$

Definición: sean X e Y variables aleatorias sobre el mismo espacio muestral S entonces diremos que X e Y son variables aleatorias independientes si:

$$P(\underbrace{[X \in C]}_A \cap \underbrace{[Y \in D]}_B) = P(X \in C) \cdot P(Y \in D) \quad \forall C, D \subset \mathbb{R}$$

Propiedad

Sean X e Y variables aleatorias sobre el mismo espacio muestral S entonces:

$$F(x, y) = P([X \leq x] \cap [Y \leq y]) = F_x(x) \cdot F_y(y) \quad \forall x, y \in \mathbb{R}$$

Con $F_x(x)$ y $F_y(y)$ las funciones de distribución acumuladas de X e Y respectivamente.

Corolario: Si X e Y son independientes

- Si X e Y son variables aleatorias discretas con función de distribución de probabilidad conjunta $P(x, y)$ entonces:

$$P(x, y) = P_x(x) \cdot P_y(y) \quad \forall x, y \in \mathbb{R}$$

- Si X e Y son variables aleatorias continuas con función de densidad de probabilidad conjunta $f(x, y)$ entonces:

$$f(x, y) = f_x(x) \cdot f_y(y) \quad \forall x, y \in \mathbb{R}$$

Volviendo al ejemplo de los contratos

X \ Y	0	1	2
0	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{9}$
1	$\frac{2}{9}$	$\frac{2}{9}$	0
2	$\frac{1}{9}$	0	0

$$P(2, 1) = 0 \text{ y } P_x(2) * P_y(1) = \frac{1}{9} * \frac{4}{9} = \frac{4}{81}$$

$\therefore X$ e Y no son variables aleatorias independientes.

En el ejemplo de los autos y camiones tampoco se cumple la independencia ya que: $0 = f(0, 0) \neq f_x(0) * f_y(0) = \frac{6}{25}$

Definiciones

Sean X_1, \dots, X_n variables aleatorias sobre S entonces:

1. Llamaremos función de probabilidad conjunta de (X_1, \dots, X_n) a

$$P(x_1, \dots, x_n) = P([X_1 = x_1] \cap \dots \cap [X_n = x_n])$$

$\forall x_i \in \mathbb{R} : 1 \leq i \leq n$ cuando las variables aleatorias son discretas.

2. Llamaremos función de densidad conjunta de (X_1, \dots, X_n) a

$$P(x_1, \dots, x_n) = \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_1 \dots dx_n$$

$\forall a_i < b_i \in \mathbb{R} : 1 \leq i \leq n$ cuando las variables aleatorias son continuas.

3. Diremos que X_1, \dots, X_n son independientes si y sólo si

$$P\left(\bigcap_{i \in I} (x_i \in [a_i, b_i])\right) = \prod_{i \in I} P(x_i \in [a_i, b_i]) \forall a_i < b_i \in \mathbb{R}$$

con $1 \leq i \leq n$ e I cualquier subconjunto de índices $\{1, \dots, n\}$.

Proposición

Sean X e Y variables aleatorias sobre un mismo espacio muestral S, con función de distribución de probabilidad o función de densidad conjunta $p(x, y)$ ó $f(x, y)$ respectivamente y sea $h : \mathbb{R}^2 \rightarrow \mathbb{R}$

Entonces:

$$E(h(x, y)) = \begin{cases} \sum_X \sum_Y h(x, y) \cdot p(x, y) & \text{si X e Y son discretas.} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) \cdot f(x, y) dx dy & \text{si X e Y son continuas.} \end{cases}$$

Esperanza de una combinación lineal de variables aleatorias

$$E(a.X + b.Y) = a.E(X) + b.E(Y) \quad \forall a, b \in \mathbb{R}$$

Demostración:

$$\begin{aligned} E(a.X + b.Y) &= \sum_X \sum_Y (ax + by) \cdot P(x, y) \\ &= a \sum_X x \sum_Y P(x, y) + b \sum_X \sum_Y y \cdot P(x, y) \\ &= a \sum_X x \cdot P_x(x) + \sum_Y y \cdot P_y(y) \\ &= a.E(x) + b.E(y) \end{aligned}$$

5.4. Covarianza

Sean X e Y variables aleatorias sobre S entonces llamamos covarianza entre X e Y , $Cov(X, Y)$ al valor:

$$Cov(x, y) = E((x - \mu_x) \cdot (y - \mu_y)) : \mu_x = E(x) \text{ y } \mu_y = E(y)$$

Proposición

$$Cov(x, y) = E(x \cdot y) - \mu_x \cdot \mu_y : \mu_x = E(x) \text{ y } \mu_y = E(y)$$

Demostración con X e Y variables aleatorias discretas

$$\begin{aligned} Cov(x, y) &= \sum_X \sum_Y \underbrace{(x - \mu_x) \cdot (y - \mu_y)}_{h(x, y)} \cdot p(x, y) \\ &= \underbrace{\sum_X \sum_Y x \cdot y \cdot P(x, y)}_{=E(x \cdot y)} - \mu_y \underbrace{\sum_X x \sum_Y P(x, y)}_{\substack{P_x(x) \\ E(x) = \mu_x}} \\ &\quad - \mu_x \underbrace{\sum_Y y \sum_x P(x, y)}_{\substack{P_y(y) \\ E(y) = \mu_y}} + \mu_x \cdot \mu_y \underbrace{\sum_X \sum_Y P(x, y)}_{=1} \\ &= E(x \cdot y) - \mu_y \mu_x - \mu_x \mu_y + \mu_x \mu_y \\ &= E(x \cdot y) - \mu_x \mu_y \end{aligned}$$

5.4.1. Propiedades de la covarianza

1. $Cov(x, y) = E(x \cdot y) - \mu_x \cdot \mu_y$
2. $Cov(x, x) = V(x)$
3. $Cov(a \cdot x + b, c \cdot y + d) = a \cdot c \cdot Cov(x, y) \quad \forall a, b, c, d \in \mathbb{R}$
4. $V(a \cdot x + b \cdot y) = a^2 \cdot V(x) + b^2 \cdot V(y) + 2 \cdot a \cdot b \cdot Cov(x, y)$

Demostración 3:

$$\begin{aligned} Cov(a.x + b, c.y + d) &= E((a.x + b)(c.y + d)) - E(a.x + b).E(c.y + d) \\ &= E(a.c.x.y + a.d.x + b.c.y + b.d) - (a.\mu_x + b).(c.\mu_y + d) \\ &= a.c.E(x.y) + a.d.\mu_x + b.c.\mu_y + b.d \\ &\quad - a.c.\mu_x.\mu_y - a.d.\mu_x - b.c.\mu_y - b.d \\ &= a.c[E(x.y) - \mu_x.\mu_y] \\ &= a.c.Cov(x, y) \end{aligned}$$

Demostración 4:

$$\begin{aligned} V(a.x + b.y) &= E((a.x + b.y)^2) - E(a.x + b.y)^2 \\ &= E(a^2.x^2 + b^2.y^2 + 2.a.b.x.y) - (a.\mu_x + b.\mu_y)^2 \\ &= a^2.E(x^2) + b^2.E(y^2) + 2.a.b.E(x.y) - [a^2.\mu_x^2 + b^2.\mu_y^2 + 2.a.b.\mu_x.\mu_y] \\ &= a^2.V(x) + b^2.V(y) + 2.a.b[E(x.y) - \mu_x.\mu_y] \\ &= a^2.V(x) + b^2.V(y) + Cov(x, y) \end{aligned}$$

5.4.2. Probabilidad conjunta y covarianza

Proposición

Sean X e Y variables aleatorias sobre S independientes entonces:

$$E(x.y) = E(x).E(y)$$

Demostración:

$$\begin{aligned} E(x.y) &= \sum_X \sum_Y (x.y).P(x, y) \\ &\stackrel{ind}{=} \sum_X \sum_Y x.y.P_x(x).P_y(y) \\ &= \sum_X x.P_x(x) \sum_Y y.P_y(y) \\ &= E(x).E(y) \end{aligned}$$

Corolario

Si X e Y Son independientes $\Rightarrow Cov(x, y) = 0$ ⁶

Corolario

Si X e Y son independientes: $V(a.x + b.y) = a^2.V(x) + b^2.V(y) \forall a, b \in \mathbb{R}$

Ejemplo

Sean X e Y variables aleatorias discretas sobre S con función de probabilidad conjunta dada por:

X \ Y	-1	0	1	$P_x(x)$
-1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
0	$\frac{1}{8}$	0	$\frac{1}{8}$	$\frac{2}{8}$
1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$
$P_y(y)$	$\frac{3}{8}$	$\frac{2}{8}$	$\frac{3}{8}$	1

$$E(x) = \sum_{-1}^1 x.P_x(x) = -\frac{3}{8} + \frac{3}{8} = 0$$

$\therefore \mu_x = \mu_y = 0$ ¿Puedo decir que X e Y son independientes?

$$P(1, 1) = \frac{1}{8} \neq \left(\frac{3}{8}\right)^2 = P_x(1).P_y(1) \therefore \text{No son independientes}$$

$$\begin{aligned} E(x.y) &= \sum_{x=-1}^1 \sum_{y=-1}^1 x.y.P(x, y) \\ &= P(-1, -1) - P(-1, 1) - P(1, -1) + P(1, 1) \\ &= 0 \end{aligned}$$

Luego X e Y son independientes $\Rightarrow Cov(x, y) = 0$ pero la recíproca puede no valer.

5.5. Definición de correlación

Sean X e Y variables aleatorias sobre S entonces se define la correlación entre X e Y como:

$$\rho(x, y) = Corr(x, y) = \frac{Cov(x, y)}{\sigma_x \cdot \sigma_y} : \sigma_x \text{ es el desvío estándar de } x.$$

⁶(Puede no valer \Leftarrow)

Proposición

Sean X e Y variables aleatorias sobre S entonces:

1. $Corr(x, x) = 1$
2. $Corr(a.x + b, c.y + d) = Sg(a.c).Corr(x, y)$
3. $\rho(x, y) \leq 1$
4. Si $y = a.x + b \Rightarrow Corr(x, y) = Sg(a)$

Demostración 2:

$$Corr(a.x + b, c.y + d) = \frac{Cov(a.x + b, c.y + d)}{\sigma_{a.x+b} \cdot \sigma_{c.y+d}} \quad (47)$$

$$= \frac{a.c.Cov(x, y)}{|a| \cdot |c| \cdot \sigma_x \cdot \sigma_y} \quad (48)$$

$$= \frac{a.c}{|a||c|} \cdot \left(\frac{Cov(x, y)}{\sigma_x \cdot \sigma_y} \right) \quad (49)$$

$$= Sg(a.c).Corr(x, y) \quad (50)$$

(47) Por definición

(48) Vale ya que si $V(a.x + b) = a^2.V(x) \Rightarrow \sigma_{ax+b} = |a| \cdot \sigma_x$

(49) $\left(\frac{Cov(x,y)}{\sigma_x \cdot \sigma_y} \right) = \rho(x, y)$

(50) $Sg(x) = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{si } x < 0 \end{cases}$

Demostración 3:

Sea $Z = t.X + y \ \forall t \in \mathbb{R}$

$$\begin{aligned} 0 &\leq V(Z) = E(z - \mu_z)^2 = E((tx + y) - (t\mu_x + \mu_y))^2 \\ 0 &\leq E((t - \mu_x) + (y - \mu_y))^2 \\ &= E[t^2(x - \mu_x)^2 + (y - \mu_y)^2 + 2.t(x - \mu_x)(y - \mu_y)] \\ &= t^2 \underbrace{E(x - \mu_x)^2}_{\sigma_x^2} + \underbrace{E(y - \mu_y)^2}_{\sigma_y^2} + 2t \underbrace{E(x - \mu_x)(y - \mu_y)}_{Cov(x,y)} \\ 0 &\leq t^2 \sigma_x^2 + \sigma_y^2 + 2t.Cov(x, y) \ \forall t \in \mathbb{R} \end{aligned}$$

Obtengo un polinomio de segundo grado en t donde:

$$a = \sigma_x^2 \qquad b = 2.Cov(x, y) \qquad c = \sigma_y^2$$

$$\begin{aligned} \text{Luego } 0 &\geq b^2 - 4.a.c = 4.Cov(x, y)^2 - 4.\sigma_x^2.\sigma_y^2 \\ \Leftrightarrow Cov(x, y)^2 - \sigma_x^2.\sigma_y^2 &\leq 0 \Rightarrow \rho(x, y)^2 \leq 1 \Rightarrow |\rho(x, y)| \leq 1 \end{aligned}$$

Demostración 4:

$$\text{Si } y = a.x + b : a \neq 0 : a, b \in \mathbb{R} \Rightarrow Corr(x, y) \stackrel{?}{=} Sg(a)$$

$$Corr(x, y) = \frac{Cov(x, ax+b)}{\sigma_x.\sigma_{ax+b}} = \frac{a}{|a|} \underbrace{\left(\frac{Cov(x, x)}{\sigma_x.\sigma_x} \right)}_{=1} = Sg(a) \quad \blacksquare$$

5.5.1. Generalización

Sean X_1, \dots, X_n Variables aleatorias y $\{a_i\}_{i=1}^n \subset \mathbb{R}$ entonces:

1. $E\left(\sum_{i=1}^n a_i.x_i\right) = \sum_{i=1}^n a_i.E(x_i)$
2. $V\left(\sum_{i=1}^n a_i.x_i\right) = \sum_{i=1}^n a_i^2.V(x_i) + 2 \sum_{i < j} a_i.a_j.Cov(x_i, x_j)$

Demostración 1

Prueba por inducción en la cantidad de términos de la sumatoria.

Si $n = 2$ ya está probado.

$$\text{Supongamos que } E\left(\sum_{i=1}^n a_i.x_i\right) = \left(\sum_{i=1}^n a_i.E(x_i)\right)$$

Veamos para $n + 1$:

$$\begin{aligned} E\left(\sum_{i=1}^{n+1} a_i.x_i\right) &= E\left(\sum_{i=1}^n a_i.x_i + (a_{n+1}.x_{n+1})\right) \\ &= E\left(\sum_{i=1}^n a_i.x_i\right) + a_{n+1}.E(x_{n+1}) \\ H.I. &= \left(\sum_{i=1}^{n+1} a_i.E(x_i)\right) \end{aligned}$$

Demostración 2

$$\begin{aligned}
 V\left(\sum_{i=1}^n a_i \cdot x_i\right) &\stackrel{Def}{=} Cov\left(\sum_{i=1}^n a_i \cdot x_i, \sum_{j=1}^n a_j \cdot x_j\right) \\
 &= \sum_{i=1}^n \sum_{j=1}^n a_i \cdot a_j \cdot Cov(x_i, x_j) \\
 &= \sum_{i=1}^n a_i^2 \cdot V(x_i) + \sum_{(i,j): i \neq j}^n a_i \cdot a_j \cdot Cov(x_i, x_j) \quad (51) \\
 &= \sum_{i=1}^n a_i^2 \cdot V(x_i) + 2 \cdot \sum_{i < j: (i,j)}^n a_i \cdot a_j \cdot Cov(x_i, x_j)
 \end{aligned}$$

(51) Luego si $i = j \Rightarrow Cov(x_i, x_j) = V(x_i)$

(51) Luego si $i \neq j \Rightarrow Cov(x_i, x_j) = Cov(x_j, x_i)$

Corolario: Si los X_i son independientes $\forall i : 1 \leq i \leq n$ se cumple que:

$$V\left(\sum_{i=1}^n a_i \cdot x_i\right) = \sum_{i=1}^n a_i^2 \cdot V(x_i)$$

Proposición

Sean X_1, \dots, X_n variables aleatorias independientes con distribución normal y $\{a_i\}_{i=1}^n \subset \mathbb{R}$ entonces:

$$\sum_{i=1}^n a_i \cdot x_i \sim N(\mu, \sigma^2) \text{ donde } \mu = \sum_{i=1}^n a_i \cdot x_i \text{ y } \sigma^2 = \sum_{i=1}^n a_i^2 \cdot V(x_i)$$

Consecuencia: Si X_1, \dots, X_n son variables aleatorias independientes e idénticamente distribuidas ⁷, con distribución normal.

Id est: con $E(x_i) = \mu$ y $V(x_i) = \sigma^2 \forall i : 1 \leq i \leq n$ entonces:

- $\bar{X} = \sum_{i=1}^n \left(\frac{1}{n}\right) \cdot E(x_i) \sim N\left(\mu, \frac{\sigma^2}{n}\right)$
- $\mu_{\bar{x}} = \sum_{i=1}^n \left(\frac{1}{n}\right) \cdot E(x_i) = \frac{1}{n} \cdot (n \cdot \mu) = \mu$
- $\sigma_{\bar{x}}^2 = \sum_{i=1}^n \left(\frac{1}{n}\right)^2 \cdot \underbrace{V(x_i)}_{=\sigma^2} = n \cdot \left(\frac{\sigma^2}{n^2}\right) = \frac{\sigma^2}{n}$

⁷Definimos esta condición como una muestra aleatoria: M.A.

Ejemplo

Sean X e Y variables aleatorias independientes e idéncicamente distribuidas $\exp(1)$.

Se pide hallar la distribución de $W = X + Y$

$$\text{Si } X \sim \exp(\lambda) \Rightarrow f_x(x) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x} & \text{si } x > 0 \\ 0 & \text{en caso contrario} \end{cases}$$

En este caso $\lambda = 1$

$$f_x(x) = \begin{cases} e^{-x} & \text{si } x > 0 \\ 0 & \text{en caso contrario} \end{cases}$$

Como son independientes $f(x, y) = f_x(x) \cdot f_y(y)$, id est:

$$f(x, y) = \begin{cases} e^{-(x+y)} & \text{si } x, y > 0 \\ 0 & \text{en caso contrario} \end{cases}$$

Luego $F_w(w) = P(W \leq w) = P(x + y \leq w) = \int \int_A f(x, y) dx dy$

Donde $A = \{(x, y) : x + y \leq w\} = \{(x, y) : y \leq w - x\}$

Si $w < 0 \Rightarrow F_w(w) = \int \int_A \underbrace{f(x, y)}_{=0} dx dy = 0$

Si $w > 0 \Rightarrow F_w(w) = \int \int_A f(x, y) dx dy \stackrel{\text{indep}}{=} \int_0^w \int_0^{w-x} e^{-(x+y)} dy dx$

$$\begin{aligned} F_w(w) &= \int_0^w e^{-x} \left(-e^{-y} \Big|_0^{w-x} \right) dx \\ &= \int_0^w e^{-x} \left(-e^{-(w-x)} + 1 \right) dx \\ &= - \int_0^w e^{-x-w+x} dx + \int_0^w e^{-x} dx \\ &= -e^{-w} (x \Big|_0^w) - (e^{-x} \Big|_0^w) \\ &= -w \cdot e^{-w} - e^{-w} + 1 \quad \forall w > 0 \end{aligned}$$

Derivo y obtengo:

$$\begin{aligned} F'_w(w) &= f_w(w) \\ &= -(e^{-w} - e^{-w}.w) + e^{-w} \\ &= w.e^{-w} \quad \forall w > 0 \text{ y } 0 \text{ en caso contrario.} \end{aligned}$$

$$\therefore f_w(w) = \begin{cases} w.e^{-w} & \text{si } w > 0 \\ 0 & \text{en caso contrario} \end{cases} \Rightarrow W \sim \Gamma(2, 1)$$

Conclusión: la distribución de la suma de variables aleatorias exponenciales no necesariamente es la distribución de una exponencial.

Ejemplo

Sean X_1, \dots, X_n una M.A. $\sim \text{Poisson}(\lambda)$ entonces:

$$X_1 + \dots + X_n \sim P(\lambda_1 + \dots + \lambda_n)$$

Demostración

Prueba por inducción en la cantidad de variables aleatorias “n”.

n = 2 : sean X e Y variables aleatorias Poisson de parámetros λ_1 y λ_2 respectivamente.

Queremos hallar la distribución de X+Y

$$X \sim P(\lambda_1) \Rightarrow P_x(x) = P[X = x] = e^{-\lambda_1} \frac{\lambda_1^x}{x!} : x \in \{0, 1, 2, \dots\} = \mathbb{Z}^{\geq 0}$$

$$Y \sim P(\lambda_2) \Rightarrow P_y(y) = P[Y = y] = e^{-\lambda_2} \frac{\lambda_2^y}{y!} : y \in \{0, 1, 2, \dots\} = \mathbb{Z}^{\geq 0}$$

Como X e Y son independientes entonces:

$$\begin{aligned} P(x, y) &= P_x(x).P_y(y) \\ &= e^{-(\lambda_1 + \lambda_2)} \frac{\lambda_1^x \lambda_2^y}{x!y!} \quad \forall x, y \in \mathbb{Z}^{\geq 0} \end{aligned}$$

Sea $W = X + Y$

Si $w \notin \mathbb{Z}^{\geq 0} \Rightarrow P[W = w] = 0$

Si $w \in \mathbb{Z}^{\geq 0}$

$$\begin{aligned}
P[W = w] &= P[X + Y = w] \\
&= P\left(\bigcup_{x=0}^w ([X = x] \cap [Y = w - x])\right) \\
&= \sum_{x=0}^w P(x, w - x) \tag{52}
\end{aligned}$$

$$\begin{aligned}
&= e^{-(\lambda_1 + \lambda_2)} \sum_{x=0}^w \frac{\lambda_1^x \lambda_2^{w-x}}{x!(w-x)!} \\
&= \frac{e^{-(\lambda_1 + \lambda_2)}}{w!} \sum_{x=0}^w \frac{w! \lambda_1^x \lambda_2^{w-x}}{x!(w-x)!} \tag{53}
\end{aligned}$$

$$\begin{aligned}
&= \frac{e^{-(\lambda_1 + \lambda_2)}}{w!} \sum_{x=0}^w \left(\frac{w!}{x!(w-x)!} \right) \lambda_1^x \lambda_2^{w-x} \\
&= \frac{e^{-(\lambda_1 + \lambda_2)}}{w!} \sum_{x=0}^w \binom{w}{x} \lambda_1^x \lambda_2^{w-x} \\
&= \frac{e^{-(\lambda_1 + \lambda_2)} \cdot (\lambda_1 + \lambda_2)^w}{w!} \tag{54}
\end{aligned}$$

(52) Unión disjunta.

(53) Multiplico y divido por $w!$.

(54) $\sum_{x=0}^w \binom{w}{x} \lambda_1^x \lambda_2^{w-x} = (\lambda_1 + \lambda_2)^w$.

$$\therefore P[W = w] = \frac{e^{-(\lambda_1 + \lambda_2)} \cdot (\lambda_1 + \lambda_2)^w}{w!} \quad \forall w \in \mathbb{Z}^{\geq 0} \Rightarrow W \sim P(\lambda_1 + \lambda_2)$$

[H.I] Supongamos que $X_1 + \dots + X_n \sim P\left(\sum_{i=1}^n \lambda_i\right)$

$$\begin{aligned}
\text{Luego } X_1 + \dots + X_n + X_{n+1} &= (X_1 + \dots + X_n) + X_{n+1} \\
&\sim P\left(\left(\sum_{i=1}^n \lambda_i\right) + \lambda_{n+1}\right) \\
&= P\left(\sum_{i=1}^{n+1} \lambda_i\right) \tag{55}
\end{aligned}$$

(55) Por H.I. tengo la suma de dos Poisson (caso base). ■

5.6. Teorema Central del Límite T.C.L

Sea X_1, \dots, X_n una M.A. con $E(x) = \mu$ y $V(x) = \sigma^2$ entonces:

$$P(\bar{X} \leq w) \approx \Phi\left(\frac{w - \mu}{\sigma} \sqrt{n}\right) \forall w$$

Si n es suficientemente grande $\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Leftrightarrow \left(\frac{\bar{x} - \mu}{\sigma}\right) \sqrt{n} \sim N(0, 1)$

La aproximación será razonable si $n \geq 30$

Observación: las variables aleatorias pueden ser discretas o continuas.

5.7. Aproximación normal a una binomial

Sea X_1, \dots, X_n una muestra aleatoria $Bernoulli(p) : p = P(E)$

$$X_i = \begin{cases} 1 & \text{si es éxito (E)} \\ 0 & \text{si es fracaso (F)} \end{cases}$$

$$E(x_i) = p \qquad E(x_i^2) = p \qquad V(x_i) = p \cdot q$$

Ya que son n ensayos independientes $\sum_{i=1}^n x_i \sim B(n, p)$

$$\text{Llamo } \hat{p} = \frac{\sum_{i=1}^n x_i}{n} \sim N\left(p, \frac{p \cdot q}{n}\right) \xrightarrow{T.C.L} \frac{\hat{p} - p}{\sqrt{\frac{p \cdot q}{n}}} \sim N(0, 1)$$

$$\frac{\hat{p} - p}{\sqrt{\frac{p \cdot q}{n}}} = \frac{\frac{1}{n} \sum_{i=1}^n x_i - n \cdot p}{\sqrt{\frac{p \cdot q}{n}}} = \frac{\sum_{i=1}^n x_i - n \cdot p}{\sqrt{n \cdot p \cdot q}} \rightarrow B(n, p) \approx X \sim N(n \cdot p, n \cdot p \cdot q)$$

$$\text{Luego la } P[X \leq x] = P\left(Z \leq \frac{x - n \cdot p}{\sqrt{n \cdot p \cdot q}}\right) \approx \Phi\left(\frac{x - n \cdot p}{\sqrt{n \cdot p \cdot q}}\right)$$

Si el n no es muy grande suele utilizarse una aproximación con conexión por continuidad:

$$\text{Luego } P(X \leq x) = P(X \leq x + 0,5) \approx \Phi\left(\frac{x + 0,5 - n \cdot p}{\sqrt{n \cdot p \cdot q}}\right)$$

Ejemplo

Supongamos que una máquina produce el 10 % de artículos defectuosos.

Se procede a detener el funcionamiento de la máquina si por lo menos el 15 % son defectuosos.

Se toma una muestra aleatoria de 100 artículos de la producción diaria.

¿Cuál es la probabilidad que deba detener la máquina para ser reparada?

Sea X = “número de artículos defectuosos en la muestra”

Luego $X \sim B(100, \frac{1}{10})$

$$\begin{aligned} P(X \geq 15) &= 1 - P(X \leq 14) \\ &\approx 1 - \Phi\left(\frac{14 - n.p}{\sqrt{n.p.q}}\right) \\ &= 1 - \Phi\left(\frac{14 - 10}{3}\right) \\ &= 1 - \Phi(1,33) \\ &= 1 - 0,9082 \\ &= 0,0918 \end{aligned}$$

Ejemplo

Sabemos que el tiempo de espera del colectivo tiene una distribución uniforme por la mañana en el intervalo (0,4) y por la tarde (0,8) minutos y que estos tiempos son independientes.

- ¿Cuál es la probabilidad de que el tiempo total de espera en 40 días sea de por lo menos tres horas y media?

Sea $W_i = X_i + Y_i : 1 \leq i \leq 40$ me pide $P\left(\sum_{i=1}^N w_i \geq 210\right) \stackrel{?}{=}$

$$\text{Si } X \sim U(a, b) \Rightarrow E(x) = \frac{a+b}{2} \text{ y la } V(x) = \frac{(b-a)^2}{12}$$

Luego la $E(W_i) = E(X_i) + E(Y_i) = 2 + 4 = 6 = \mu$

Y la $V(W_i) = V(X_i) + V(Y_i) = \frac{16}{12} + \frac{64}{12} = \frac{80}{12} = \frac{20}{3} = \sigma^2$

$$\therefore \overline{W} \sim N\left(6, \frac{20/3}{40}\right) = N\left(6, \frac{1}{6}\right)$$

- ¿Cuántos días deberíamos tomar de forma tal de asegurar con una probabilidad mayor al 99 % para garantizar que el tiempo total de espera sea de por lo menos tres horas y media ?

$$0,99 \stackrel{exijo}{<} P\left(\sum_{i=1}^n w_i \geq 210\right)$$

Para obtener el promedio dividido por n la desigualdad y obtengo:

$$0,99 < P\left(\sum_{i=1}^n \frac{w_i}{n} \geq \frac{210}{n}\right)$$

$$\text{Luego } \bar{W} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$P\left(\left(\frac{\bar{W}-\mu}{\sigma}\right) \cdot \sqrt{n} \geq \left(\frac{\frac{210}{n}-\mu}{\sigma}\right) \sqrt{n}\right)$$

$$P\left(\left(\frac{\bar{W}-\mu}{\sigma}\right) \cdot \sqrt{n} \geq \left(\frac{210-n \cdot \mu}{\sqrt{n} \cdot \sigma}\right)\right)$$

$$0,99 < P\left(Z \geq \left(\frac{210-n \cdot \mu}{\sqrt{n} \cdot \sigma}\right)\right)$$

$$0,99 < 1 - \Phi\left(\frac{210-n \cdot \mu}{\sqrt{n} \cdot \sigma}\right)$$

$$\Phi\left(\frac{210-n \cdot \mu}{\sqrt{n} \cdot \sigma}\right) < 0,01$$

$$\text{Llamando a } z = \left(\frac{210-n \cdot \mu}{\sqrt{n} \cdot \sigma}\right)$$

$$\text{Obtengo que : } 1 - \Phi(z) < 0,01 \Leftrightarrow \Phi(z) \geq 0,99 \Leftrightarrow z \geq 2,33$$

$$\therefore -z = \frac{210-n \cdot \mu}{\sqrt{n} \cdot \sigma} \leq -2,33 \Leftrightarrow \frac{n \cdot \mu - 210}{\sqrt{n} \cdot \sigma} \geq 2,33$$

Reemplazando μ y σ por sus valores me queda un polinomio cuadrático en n donde las raíces son: $x_1 = 29,55$ y $x_2 = 41,66$.

Respuesta: con tomar $n \geq 42$ días alcanzará. ■

Ejemplo

El tiempo que tarda un empleado en procesar el pedido de cada cliente es una variable aleatoria con media $1,5min$ y una desviación estándar de 1 minuto.

Suponiendo que los tiempos que tarda en procesar n pedidos son independientes. Se pide:

1. ¿Cuál es la probabilidad aproximada de que se puedan procesar los pedidos de 100 clientes en menos de dos horas?

Pasando en limpio tengo: $\mu_x = 1,5'$ y $\sigma_x = 1'$ con $n = 100$

$$P\left(\sum_{i=1}^n \frac{x_i}{n} \leq \frac{120'}{n}\right) = P(\bar{X} \leq 1,20) : \bar{X} \sim N\left(\mu, \frac{\sigma_x^2}{n}\right)$$

Estandarizo y obtengo:

$$P\left(Z \leq \frac{1,20-1,5}{\sqrt{\frac{1}{100}}}\right) = \Phi(-3) = 1 - \Phi(3) = 1 - 0,9987 = 0,0013$$

2. Determinar el menor valor t_0 tal que con una probabilidad de por lo menos 0,9 se puedan procesar 100 pedidos en un tiempo de a lo sumo t_0 .

$$0,9 \stackrel{\text{exijo}}{\leq} P(\bar{X} \leq \frac{t_0}{n})$$

Primero estandarizo y obtengo:

$$P\left(\left(\frac{\bar{X} - \mu}{\sigma}\right) \cdot \sqrt{n} \leq \left[\frac{\frac{t_0}{n} - 1,5}{1}\right] \cdot \sqrt{100}\right)$$

Llamando a: $z = \left(\frac{t_0}{100} - 1,5\right) \cdot 10$

Obtengo que: $\Phi(z) \stackrel{\text{exijo}}{\geq} 0,9 \Rightarrow z \geq 1,29$

$$\begin{aligned}\left(\frac{t_0}{100} - 1,5\right) \cdot 10 &\geq 1,29 \\ t_0 &\geq \left[\frac{1,29}{10} + 1,5\right] \cdot 100 \\ t_0 &\geq 162,9\end{aligned}$$

Ejemplo

Hallar la probabilidad aproximada para una variable aleatoria Poisson de parámetro 50 de tomar valores entre 35 y 70.

$$X \sim P(50) \Rightarrow P(35 \leq X \leq 70) \stackrel{?}{=}$$

Truco para usar el T.C.L.

Sean X_1, \dots, X_n una muestra aleatoria poisson de parámetro $\lambda = 1 \Rightarrow$

$\sum_{i=1}^N X_i \approx P(50)$. Luego por T.C.L $\frac{\sum_{i=1}^N X_i}{n} \sim N\left(1, \frac{1}{n}\right)$ Donde $n = 50$

$$\begin{aligned} P(35 \leq X \leq 70) &= P\left(\frac{35}{50} \leq \frac{\sum_{i=1}^{50} X_i}{50} \leq \frac{70}{50}\right) \\ &= P\left(\frac{\frac{35}{50} - 1}{\frac{1}{\sqrt{50}}} \leq \frac{\bar{X} - 1}{\frac{1}{\sqrt{50}}} \leq \frac{\frac{70}{50} - 1}{\frac{1}{\sqrt{50}}}\right) \\ &= P(-2,12 \leq z \leq 2,83) \\ &= \Phi(2,83) - \Phi(-2,12) = \Phi(2,83) - (1 - \Phi(2,12)) \\ &= 0,977 - 1 + 0,9830 = 0,9807 \end{aligned}$$

■

6. Estimación puntual

Definición: sea X_1, \dots, X_n una muestra aleatoria con distribución acumulada que dependa de θ .

Se dice que $\hat{\theta} = h(x_1, \dots, x_n)$ es un estimador para θ si $\hat{\theta}$ es una variable aleatoria y $\hat{\theta} = \theta + \epsilon$.

Definición: se dice que un estimador $\hat{\theta}$ es **insesgado** para θ si: $E(\hat{\theta}) = \theta$

Proposición

Sea X_1, \dots, X_n una muestra aleatoria : $E(x) = \mu$ y la $V(x) = \sigma^2$ entonces:

- $\hat{\theta}_1 = \bar{X}$ es insesgado para μ
- $\hat{\theta}_2 = S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$ es insesgado para σ^2

Demostración:

(a) Si $\hat{\theta}_1 = \bar{X}$ es insesgado para $\mu \Rightarrow E(\hat{\theta}_1) = \mu$

$$E(\hat{\theta}_1) = E(\bar{X}) = E\left(\sum_{i=1}^n \frac{x_i}{n}\right) = \sum_{i=1}^n \frac{1}{n} \cdot E(x_i) = \sum_{i=1}^n \frac{1}{n} \cdot \mu = \mu$$

$$\begin{aligned} (b) \hat{\theta}_2 = S^2 &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{1}{n-1} \sum_{i=1}^n ((x_i - \mu) - (\bar{x} - \mu))^2 \\ &= \frac{1}{n-1} \left[\sum_{i=1}^n (x_i - \mu)^2 + \sum_{i=1}^n (\bar{x} - \mu)^2 - 2 \cdot n \cdot (\bar{x} - \mu) \cdot \sum_{i=1}^n \frac{x_i - \mu}{n} \right] \\ &= \frac{1}{n-1} \left[\sum_{i=1}^n (x_i - \mu)^2 + n \cdot (\bar{x} - \mu)^2 - 2 \cdot n \cdot (\bar{x} - \mu)^2 \right] \\ &= \frac{1}{n-1} \left[\sum_{i=1}^n (x_i - \mu)^2 - n \cdot (\bar{x} - \mu)^2 \right] \end{aligned}$$

$$\begin{aligned} \therefore E(\hat{\theta}_2) &= \frac{1}{n-1} \left[\sum_{i=1}^n \underbrace{E(x_i - \mu)^2}_{V(x_i = \sigma^2)} - n \cdot \underbrace{E(\bar{x} - \mu)^2}_{V(\bar{x}) = \frac{\sigma^2}{n}} \right] \\ &= \frac{1}{n-1} \cdot [n \cdot \sigma^2 - \sigma^2] = \frac{1}{n-1} \cdot (n-1) \cdot \sigma^2 = \sigma^2 \end{aligned}$$

■

Criterio de elección

Supongamos $\hat{\theta}_1$ y $\hat{\theta}_2$ son estimadores insesgados para θ debemos elegir el de menor varianza.

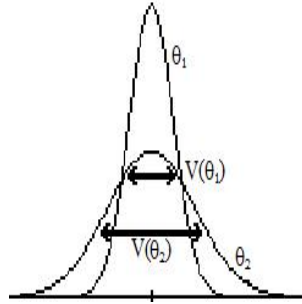


Figura 4: $V(\hat{\theta}_1) < V(\hat{\theta}_2) \therefore$ En este caso elegimos θ_1

Ejemplo

Sea X_1, \dots, X_n una muestra aleatoria Bernoulli de parámetro p .

- $\hat{\theta} = \bar{X}$ es un estimador insesgado para $E(x) = p$ y $E(x^2) = p$ y la $V(x) = p \cdot (1 - p)$
- Sea X_1, \dots, X_n una M.A. $\sim U(0, \theta) \Rightarrow \bar{X}$ no es insesgado para θ
 $\mu = \frac{\theta}{2}$ y la $E(\bar{x}) = \frac{\theta}{2} \neq \theta$

De todas maneras si elegimos $\hat{\theta}_1 = 2 \cdot \bar{X} \Rightarrow$ si es insesgado para θ

Recordemos que si: $X \sim U(0, \theta) \Rightarrow f_x(x) = \begin{cases} \frac{1}{\theta} & \text{si } x \in [0, \theta] \\ 0 & \text{en caso contrario} \end{cases}$

$$F_x(x) = \begin{cases} 0 & \text{si } x = 0 \\ \frac{x}{\theta} & \text{si } x \in (0, \theta) \\ 1 & \text{si } x \geq \theta \end{cases} \quad \therefore F_{\hat{\theta}}(x) = \begin{cases} 0 & \text{si } x = 0 \\ (\frac{x}{\theta})^n & \text{si } x \in (0, \theta) \\ 1 & \text{si } x \geq \theta \end{cases}$$

$$f_{\hat{\theta}}(x) = F'_{\hat{\theta}}(x) = \begin{cases} \frac{n \cdot x^{n-1}}{\theta^n} & \text{si } x \in (0, \theta) \\ 0 & \text{en caso contrario} \end{cases}$$

$$\begin{aligned}
E(\hat{\theta}_2) &= \int_{-\infty}^{\infty} x \cdot f_{\hat{\theta}_2}(x) \, dx \\
&= \frac{n}{\theta^n} \int_0^{\theta} x^n \, dx \\
&= \frac{n}{\theta^n} \left(\frac{x^{n+1}}{n+1} \right) \Big|_0^{\theta} \\
&= \frac{n}{\theta^n} \left(\frac{\theta^{n+1}}{n+1} \right) \\
&= \frac{n \cdot \theta}{n+1} \neq \theta \quad \forall n \in \mathbb{N}
\end{aligned}$$

$\therefore \hat{\theta}_2$ No es insesgado para θ

Sea $\hat{\theta}_3 = \frac{N+1}{N} * \hat{\theta}_2 \Rightarrow \hat{\theta}_3$ Es un estimador insesgado para θ

Calculo la $V(\hat{\theta}_3)$ Para facilitar calculos calculamos la de $\hat{\theta}_2$ ya que es una combinación lineal de la que buscamos.

$$\begin{aligned}
E((\hat{\theta}_2)^2) &= \int_{-\infty}^{\infty} x^2 \cdot f_{\hat{\theta}_2}(x) \, dx \\
&= \frac{n}{\theta} \int_0^{\theta} x^{n+1} \, dx \\
&= \frac{n}{\theta^n} \left(\frac{x^{n+2}}{n+2} \right) \Big|_0^{\theta} \\
&= \frac{n \cdot \theta^2}{n+2}
\end{aligned}$$

$$V((\hat{\theta}_2)^2) = E((\hat{\theta}_2)^2) - E(\hat{\theta}_2)^2 = \frac{n \cdot \theta^2}{n+2} - \frac{n^2 \theta^2}{(n+1)^2} = \frac{n \cdot \theta^2}{(n+2)(n+1)^2}$$

$$\begin{aligned}
\therefore V(\hat{\theta}_3) &= V\left(\frac{n+1}{n} \cdot \hat{\theta}_2\right) \\
\left(\frac{n+1}{n}\right)^2 \cdot V(\hat{\theta}_2) &= \frac{\theta^2}{n \cdot (n+2)}
\end{aligned}$$

Calculo la $V(\hat{\theta}_1)$

$$V(\bar{x}) = \frac{\sigma^2}{n} = \frac{\frac{(b-a)^2}{12}}{n} = \frac{\theta^2}{12n}$$

Como $b - a = \theta$

Luego $V(\hat{\theta}_1) = 4.V(\bar{X}) = 4.\frac{\theta^2}{12.n} = \frac{\theta^2}{3.n}$

Ahora comparamos para saber con que estimador nos quedamos

Partimos de $n(n+2) \geq 3.n \quad \forall n \in \mathbb{N}$

Luego $V(\hat{\theta}_3) < V(\hat{\theta}_1) \quad \forall n \in \mathbb{N}$

$\therefore \hat{\theta}_3$ es mejor estimador que $\hat{\theta}_1$

Teorema

Si X_1, \dots, X_n es una muestra aleatoria con distribución $N(\mu, \sigma^2)$ entonces $\hat{\mu} = \bar{X}$ es insesgado y de mínima varianza entre todos los insesgados para μ .

$$V(\hat{\mu}) \leq V(\hat{\theta}) \quad \forall \hat{\theta} \text{ estimador de } \mu \text{ se lo denota IMVU.}$$

Sea X_1, \dots, X_n una muestra aleatoria con distribución f que depende de θ .

Llamamos **error estándar** de $\hat{\theta}$ a la $\sqrt{V(\hat{\theta})}$

Si el error estándar depende de algún parámetro desconocido entonces al estimarlo obtenemos lo que llamamos **error estimado** $\left(\widehat{\sqrt{V(\hat{\theta})}}\right)$

Volviendo al problema de una M.A. $\sim U(0, \theta)$

$$\hat{\theta}_3 = \frac{n+1}{n} \cdot \max x_i$$

Luego el error estándar sería $\sqrt{V(\hat{\theta}_3)} = \sqrt{\frac{\theta^2}{n.(n+2)}}$

Luego el error estándar estimado es : $\widehat{\sqrt{V(\hat{\theta}_3)}} = \sqrt{\frac{(\hat{\theta}_3)^2}{n.(n+2)}}$

Nota:

El estimador es una variable aleatoria \rightarrow Notación: \bar{X}

La estimación es un valor \rightarrow Notación: \bar{x}

- Sea X_1, \dots, X_n una M.A $\sim N(\mu, \sigma^2)$

El estimador para $\mu \rightarrow \bar{x}$

$$\text{Error estándar de } \bar{x} \rightarrow \sqrt{V(\bar{x})} = \sqrt{\frac{\sigma^2}{n}}$$

Si σ es desconocido \rightarrow reemplazo σ^2 por S^2

Luego el error estándar aproximado es: $\left(\frac{S}{\sqrt{n}}\right)$

6.1. Métodos para la estimación de los parámetros de la distribución

- Método de los momentos.
- Método de máxima verosimilitud.

6.1.1. Método de los Momentos

Definición: sea X_1, \dots, X_n una muestra aleatoria con distribución F que depende de parámetros $\theta_1, \dots, \theta_M$.

Llamamos k-ésimo momento poblacional a: $E(x^k)$

y k-ésimo momento muestral a: $\frac{\sum_{i=1}^n x_i^k}{n} \quad \forall k \in \mathbb{N}$

Descripción del método:

$$\text{Sistema de m ecuaciones} = \begin{cases} E(x) &= \bar{x} \\ E(x^2) &= \frac{\sum_{i=1}^n x_i^2}{n} \\ \vdots &= \vdots \\ E(x^k) &= \frac{\sum_{i=1}^n x_i^k}{n} \end{cases}$$

Luego si $\hat{\theta}_1, \dots, \hat{\theta}_M$ son soluciones del sistema de m ecuaciones diremos que son los estimadores por el método de los momentos para $\theta_1, \dots, \theta_M$

Ejemplos

1. Sea X_1, \dots, X_n una M.A. $\sim U(0, \theta)$

$$\frac{\theta}{2} = E(x) = \bar{X} \Rightarrow \hat{\theta} = 2.\bar{X}$$

2. Sea X_1, \dots, X_n una M.A. $\sim N(\mu, \sigma^2)$

con μ y σ desconocidos luego $m = 2$

$$\text{Luego } \mu = E(x) = \bar{X}$$

$$\sigma^2 + \mu^2 = E(x^2) = \frac{\sum_{i=1}^n x_i^2}{n}$$

$$\hat{\mu} = \bar{x}$$

$$\hat{\sigma} = \frac{\sum_{i=1}^n x_i^2 - n.\bar{x}^2}{n} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \frac{n-1}{n} * S^2$$

Luego $\hat{\mu}$ y $\hat{\sigma}^2$ Son E.M.M para μ y σ^2 respectivamente.

$$E(\hat{\sigma}^2) = \frac{n-1}{n} \cdot E(S^2) = \frac{n-1}{n} * \sigma^2 < \sigma^2 \therefore \text{No es insesgado para } \sigma^2$$

Pero $\hat{\mu}$ si lo es para μ

3. Sea X_1, \dots, X_n una M.A. $\sim \text{Bernoulli}(p)$

$$E(x) = \overline{X} \stackrel{\text{def}}{=} p \therefore \hat{p} = \frac{\sum_{i=1}^n x_i}{n}$$

4. Sea X_1, \dots, X_n una M.A. $\sim P(\lambda)$

$$E(x) = \overline{X} \stackrel{\text{def}}{=} \lambda$$

\therefore El estimador para el método de los momentos para $\hat{\lambda} = \overline{X}$

Id est: $E(\hat{\lambda}) = E(\overline{X}) = \lambda$ y es insesgado para λ .

Recordar:

Sea X_1, \dots, X_n donde $E(x) = \mu$ y la $V(x) = \sigma^2$ entonces

$$E(\overline{X}) = \mu \text{ y la } V(\overline{X}) = \frac{\sigma^2}{n}$$

5. Recordando que si $X \sim \exp(\lambda) \Rightarrow E(x) = \frac{1}{\lambda}$ y la $V(x) = \frac{1}{\lambda^2}$

Sea X_1, \dots, X_n una M.A. $\sim \exp(\lambda)$

$$\frac{1}{\lambda} = E(x) = \overline{X}$$

Luego $\hat{\lambda} = \frac{1}{\overline{X}}$ Es el E.M.M. para λ

$$f_x(x) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x} & \text{si } x > 0 \\ 0 & \text{en caso contrario} \end{cases}$$

$$\text{Luego } P(\overline{X} > 0) = 1$$

6.1.2. Método de Máxima Verosimilitud

Descripción del método: sea X_1, \dots, X_n una M.A. con función de densidad conjunta:

- $f(x_1, \dots, x_n, \theta_1, \dots, \theta_n)$ o función de probabilidad conjunta.
- $P(x_1, \dots, x_n, \theta_1, \dots, \theta_n)$ si es discreta.

Dependiendo si la muestra es de una variable aleatoria continua o una variable aleatoria discreta respectivamente.

Entonces $\hat{\theta}_1, \dots, \hat{\theta}_n$ son los estimadores de máxima verosimilitud si para todo $\theta_1, \dots, \theta_n$

Si para cada $\tilde{x} = (x_1, \dots, x_n)$

$$f(\tilde{x}, \hat{\theta}) \geq f(\tilde{x}, \theta) \text{ ó } P(\tilde{x}, \hat{\theta}) \geq P(\tilde{x}, \theta) \quad \forall \theta \in \mathbb{R}^n$$

Donde $\tilde{\theta} = (\theta_1, \dots, \theta_M)$ y $\hat{\tilde{\theta}} = (\hat{\theta}_1, \dots, \hat{\theta}_M)$

Ejemplos

1. Sea X_1, \dots, X_n una M.A. $\sim P(\lambda)$

recordando que $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad \forall k \in \mathbb{Z}^{\geq 0}$

$\tilde{x} = (x_1, \dots, x_n)$ Luego $\theta_1 = \lambda \therefore M = 1$

Suponer que $x_i \in \mathbb{Z}^{\geq 0} \quad \forall_i : 1 \leq i \leq N$ Para que no nos de 0 la productoria.

$$P(\tilde{x}, \lambda) \stackrel{\text{independes}}{=} \prod_{i=1}^n P(x_i, \lambda) \stackrel{\text{ident} = \text{distr}}{=} \prod_{i=1}^n e^{-\lambda} \cdot \frac{\lambda^{x_i}}{x_i!} = e^{-n \cdot \lambda} \cdot \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} = H(\lambda)$$

Luego maximizar la función $H(\lambda)$ es equivalente a maximizar el logaritmo

$$\therefore \log(h(\lambda)) = -n \cdot \lambda + \left(\sum_{i=1}^n x_i \right) \cdot \log(\lambda) - \lambda \cdot n \cdot \prod_{i=1}^n x_i!$$

Derivo respecto de λ : $\frac{d}{d\lambda} \log(H(\lambda)) = \frac{-n \cdot \lambda + \sum_{i=1}^n x_i}{\lambda} \stackrel{\text{exijo}}{=} 0$

$$-n \cdot \lambda + \sum_{i=1}^n x_i = 0 \Rightarrow \lambda = \overline{X}$$

Tengo que chequear que sea máximo es decir que la derivada segunda respecto de λ sea negativa en el punto

$$\frac{d^2}{d\lambda^2} \log(H(\lambda)) = \frac{-\sum_{i=1}^n x_i}{\lambda^2} < 0$$

$\therefore \hat{\lambda} = \bar{X}$ Es el E.M.V para λ .

2. Sea X_1, \dots, X_n una M.A. $\sim \exp(\lambda) : \lambda > 0$

Sabemos que $f_x(x) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x} & \text{si } x > 0 \\ 0 & \text{en caso contrario} \end{cases}$

Y que $E(x) = \frac{1}{\lambda}$ y la $V(x) = \frac{1}{\lambda^2}$

Sea $\tilde{x} \in \mathbb{R}^n$

$$\begin{aligned} \text{Luego } f(\tilde{x}, \lambda) &\stackrel{\text{indeptes}}{=} \prod_{i=1}^n f_x(x_i, \lambda) \\ &= \begin{cases} \lambda^n \cdot e^{-\lambda \cdot \sum_{i=1}^n x_i} & \text{si } x_i > 0 \forall_i : 1 \leq i \leq n \\ 0 & \text{en caso contrario} \end{cases} \end{aligned}$$

Si $X_i > 0 \forall_i : 1 \dots n$

$$f(\tilde{x}, \lambda) = \lambda^n \cdot e^{-\lambda \cdot \sum_{i=1}^n x_i}$$

Maximizar la función es lo mismo que maximizar el: $\log(f(\tilde{x}, \lambda))$

$$\therefore \log(f(\tilde{x}, \lambda)) = n \cdot \log(\lambda) - \sum_{i=1}^n x_i = G(\lambda)$$

$$G'(\lambda) = \frac{n}{\lambda} - \sum_{i=1}^n x_i = 0 \Leftrightarrow n - \lambda \cdot \sum_{i=1}^n x_i = 0 \Leftrightarrow \lambda = \frac{1}{\bar{X}}$$

Como es un punto de máximo obtengo que $\hat{\lambda} = \frac{1}{\bar{X}}$ es el estimador de máxima verosimilitud para λ

3. Sea X_1, \dots, X_n una M.A. $\sim N(\mu, \sigma^2)$

$$\begin{aligned} f(\underset{\sim}{x}, \underset{\sim}{\theta}) &= \prod_{i=1}^n f_x(x_i, \underset{\sim}{\theta}) \\ &= \prod_{i=1}^n \left[\frac{1}{\sqrt{2\pi \cdot \sigma^2}} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}} \right] \\ &= (2\pi \cdot \sigma^2)^{\frac{n}{2}} \cdot e^{-\frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2}} \end{aligned}$$

Por otra parte queremos reducir la siguiente expresión:

$$\begin{aligned} \star \sum_{i=1}^n (x_i - \mu)^2 &= \sum_{i=1}^n ((x_i - \bar{x}) - (\mu - \bar{x}))^2 \\ &= \sum_{i=1}^n (x_i - \mu)^2 + \sum_{i=1}^n (\mu - \bar{x})^2 - 2 \cdot (\mu - \bar{x}) \cdot \underbrace{\sum_{i=1}^n (x_i - \bar{x})}_{=0} \\ &= \sum_{i=1}^n (x_i - \bar{x})^2 + n \cdot (\mu - \bar{x})^2 \end{aligned}$$

Volviendo reemplazo \star y obtengo

$$\begin{aligned} f(\underset{\sim}{x}, \underset{\sim}{\theta}) &= (2\pi \cdot \sigma^2)^{\frac{n}{2}} \cdot e^{\frac{\sum_{i=1}^n (x_i - \bar{x})^2 + n \cdot (\mu - \bar{x})^2}{2\sigma^2}} \\ &= \underbrace{\left[(2\pi \cdot \sigma^2)^{-\frac{n}{2}} \cdot e^{-\frac{\sum x_i - \bar{x}^2}{2\sigma^2}} \right]}_{\text{función que depende de } \sigma^2} \cdot \underbrace{\left[e^{-\frac{n \cdot (\mu - \bar{x})^2}{2\sigma^2}} \right]}_{\text{depende de } \theta} \end{aligned}$$

$$\text{Acoto } \left[e^{-\frac{n \cdot (\mu - \bar{x})^2}{2\sigma^2}} \right] \leq e^0 = 1$$

Sabiendo que $\frac{n \cdot (\mu - \bar{x})^2}{2\sigma^2} = 0 \Leftrightarrow \mu = \bar{x}$

Luego acoto la función:

$$\leq \left[(2\pi \cdot \sigma^2)^{-\frac{n}{2}} \cdot e^{-\frac{\sum x_i - \bar{x}^2}{2\sigma^2}} * 1 \right] = G((\sigma)^2)$$

$$\text{Aplico } \log(G(\sigma)^2) = -\frac{n}{2} \log(2\pi\sigma^2) - \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{2\sigma^2}$$

$$\text{Luego } -\frac{n}{2} \log(2\pi\sigma^2) - \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{2\sigma^2} = 0$$

$$\Leftrightarrow -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) - \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{2\sigma^2}$$

$$\frac{d}{d\sigma^2} \log(\sigma^2) = 0 - \frac{n}{2} \frac{1}{\sigma^2} + \frac{1}{\sigma^4} \left(\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{2\sigma^2} \right)$$

$$\Leftrightarrow 0 = \frac{-n\sigma^2 + \sum_{i=1}^n (x_i - \bar{x})^2}{2\sigma^4}$$

$$\Leftrightarrow \sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \text{ Y sé que es un punto de máximo.}$$

$$\therefore \text{ El E.M.V para } \theta = (\mu, \sigma^2) \text{ es } \hat{\theta}_{\sim} = \left(\bar{x}, \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \right)$$

$$4. f(x, \theta) = \prod_{i=1}^N f_x(x_i, \theta)$$

$$\text{Donde } f_x(x) = \begin{cases} \frac{1}{\theta^n} & \text{si } x_i \in (0, \theta) \forall i : 1 \leq i \leq n \\ 0 & \text{en caso contrario} \end{cases}$$

$$f(x, \theta) = \left[\frac{1}{\theta^n} I_{(0, \infty)}(\min x_i) \cdot I_{(0, \infty)}(\max x_i) \right]$$

$$\text{Donde } I_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{en caso contrario} \end{cases}$$

$$\text{Supongamos } x_i \in (0, \theta) \forall i : 1 \leq i \leq n \Rightarrow 0 < \min_{x_i} \leq \max_{x_i} < \theta$$

$$\text{Si } \exists i : x_i \notin (0, \theta) \Rightarrow \begin{cases} x_i \leq 0 & I_{(0, \infty)}(\min x_i) = 0 \\ x_i \geq 0 & I_{(0, \infty)}(\max x_i) > 0 \end{cases}$$

$$\Rightarrow I_{(0, \infty)}(\max x_i) = 0 \text{ Para cada } x_{\sim}$$

$$f(x, \theta) = \frac{1}{\theta^n} I_{(0, \infty)}(\min x_i) * I_{(0, \infty)}(\max x_i)$$

$$\text{Por otra parte si } \theta > \max_{x_i} > 0 \Rightarrow f(x, \theta) = \frac{1}{\theta^n}$$

$$\text{Por último si } \theta \leq \max_{x_i} \Rightarrow f(x, \theta) = 0$$

$$\text{Luego el E.M.V para } \theta \text{ es } \hat{\theta}_{EMV} = \max x_i \text{ Recordar el } \hat{\theta}_{EMM} = 2\bar{X}$$

6.1.3. Propiedades de los E.M.V

1. Propiedad de invarianza

Sea X_1, \dots, X_n una muestra aleatoria con distribución $F(x, \theta) : \theta \in \mathbb{R}^n$

Si $\hat{\theta}$ es el E.M.V. para θ y sea $h : \mathbb{R}^n \rightarrow \mathbb{R}$.

Entonces el E.M.V para $h(\theta)$ es $h(\hat{\theta})$

2. Si n es suficientemente grande entonces el estimador de máxima verosimilitud de θ es asintóticamente insesgado, id est: $E(\hat{\theta}) = \theta$

Ejemplo

Sea X_1, \dots, X_n una M.A. $\sim N\left(\mu, \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}\right)$

$$\hat{\theta}_{E.M.V} = \left(\bar{x}, \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}\right)$$

Se pide hallar el E.M.V. para el percentil p de una $N(\mu, \sigma^2)$

$$\eta(p) : p = F(\eta(p)) = P(X \leq \eta(p))$$

$$\text{Estandarizo y obtengo } P\left(\underbrace{\frac{X - \mu}{\sigma}}_Z \leq \underbrace{\frac{\eta(p) - \mu}{\sigma}}_{\mathfrak{z}}\right)$$

$$\mathfrak{z} = \frac{\eta(p) - \mu}{\sigma} \therefore \eta(p) = \underbrace{\sigma \mathfrak{z} + \mu}_{h(\mu, \sigma^2)}$$

Luego por la propiedad de **invarianza** de los E.M.V. resulta que el E.M.V.

para el percentil de una $N(\mu, \sigma^2)$ es igual a $h\left(\hat{\theta}_{E.M.V.}\right) = \mathfrak{z} \cdot \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} + \bar{X}$

7. Intervalos de confianza basados en una sola muestra

Sea X_1, \dots, X_n una muestra aleatoria con distribución acumulada $F(x, \theta)$ con θ desconocido.

Objetivo: para construir un intervalo de confianza (IC) para el parámetro θ debemos obtener un estadístico $\hat{\theta} = \hat{\theta}(x_1, \dots, x_n)$ que me permita generar intervalos de longitud pequeña (precisión) que con una alta confiabilidad (nivel de confianza) contengan al parámetro θ .

7.1. Método general para obtener un IC para θ

Sea X_1, \dots, X_n una muestra aleatoria con distribución $F(x, \theta)$.

Obtener un estadístico $h(x_1, \dots, x_n, \theta)$ que dependa de θ pero su distribución no.

Seleccionar un **nivel de confianza** $(1 - \alpha)$ para el intervalo y encontrar $a < b$ tales que cumplan:

$$P(a \leq h(x_1, \dots, x_n, \theta) \leq b) = (1 - \alpha)$$

A partir de la expresión $a \leq h(x_1, \dots, x_n, \theta) \leq b$ obtener:

- $L(x_1, \dots, x_n)$ (Low) llamada cota aleatoria inferior.
- $U(x_1, \dots, x_n)$ (Upper) llamada cota aleatoria superior.

Tales que $P\left(\theta \in [L(x_1, \dots, x_n), U(x_1, \dots, x_n)]\right) = 1 - \alpha$

Ejemplo

Sea X_1, \dots, X_n una M.A. $\sim N(\mu, \sigma^2)$ con σ^2

Problema:

Hallar un IC de nivel $(1 - \alpha)$ para $\theta = \mu$

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \leftrightarrow \underbrace{\left(\frac{\bar{X} - \mu}{\sigma}\right) \sqrt{n}}_{h(x_1, \dots, x_n, \mu)} \sim N(0, 1)$$

Luego fijado un nivel $(1 - \alpha)$

$$P(a \leq h(x_1, \dots, x_n, \mu) \leq b) = 1 - \alpha$$

Elijo $a = -z$ y $b = z$

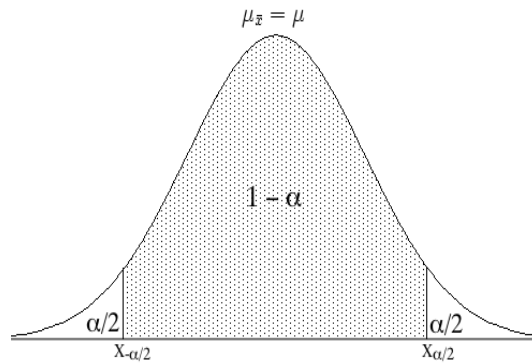


Figura 5: $b = X_{\alpha/2}$

Definición: llamaremos z_γ al valor tal que $P(Z \geq z_\gamma) = \gamma$

$$-z_{\frac{\alpha}{2}} \leq \left(\frac{\bar{X} - \mu}{\sigma} \right) \sqrt{n} \leq z_{\frac{\alpha}{2}}$$

$$-z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \bar{X} - \mu \leq z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$\underbrace{\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}}_{l(x_1, \dots, x_n)} \leq \mu \leq \underbrace{\bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}}_{U(x_1, \dots, x_n)}$$

Luego el intervalo de de confianza de nivel $(1 - \alpha)$ aleatorio para μ es:

$$\boxed{\bar{X} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}}$$

Observaciones:

- El intervalo de confianza está centrado en \bar{X} y su longitud es igual a $2 \cdot z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$
- Al aumentar el nivel de confianza $\Rightarrow z_{\frac{\alpha}{2}}$ crece .
- Luego aumentar la longitud del intervalo \Rightarrow pérdida de precisión.

Si se quiere una precisión de a lo sumo w y un nivel $(1 - \alpha)$ la única posibilidad que nos queda es elegir el tamaño de muestra adecuado.

$$\text{Long (IC)} = 2 \cdot z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq w \therefore n \geq \left(\frac{2 \cdot z_{\frac{\alpha}{2}} \cdot \sigma}{w} \right)^2$$

7.2. IC para $\theta = \mu$ con muestras suficientemente grandes

Sea X_1, \dots, X_n una muestra aleatoria con $E(x) = \mu$ y con $V(x) = \sigma^2$

Como n es suficientemente grande puedo usar el teorema central del límite que dice:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Leftrightarrow \left(\frac{\bar{X} - \mu}{\sigma}\right) \cdot \sqrt{n} \sim N(0, 1)$$

- si σ es conocido entonces:

$$H(x_1, \dots, x_n, \mu) \Rightarrow \left(\frac{\bar{X} - \mu}{\sigma}\right) \cdot \sqrt{n} \sim N(0, 1)$$

entonces fijado un nivel de confianza $(1 - \alpha)$ y pensando de manera similar a la anterior tenemos que:

$\left[l(x_1, \dots, x_n), U(x_1, \dots, x_n)\right] = \bar{X} \pm z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ es un I.C. de nivel aproximado $(1 - \alpha)$ para $\theta = \mu$

- Si σ no es conocido entonces:

Sabemos que S^2 es un estimador insesgado para σ^2 id est: $E(S^2) = \sigma^2$ y tiene la propiedad de consistencia:

$$\forall \epsilon > 0 \Rightarrow P(|S^2 - \sigma^2| \geq \epsilon) \xrightarrow{n \rightarrow \infty} 0$$

$$\text{o lo mismo } P(|S^2 - \sigma^2| \leq \epsilon) \xrightarrow{n \rightarrow \infty} 1$$

$$\text{Además } \left(\frac{\bar{X} - \mu}{S}\right) \cdot \sqrt{n} \sim N(0, 1)$$

Si $n > 40$ luego un IC de nivel $(1 - \alpha)$ aproximado para μ es: $\bar{X} \pm z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$

Ejemplo

Sea X_1, \dots, X_n una M.A. $\sim \text{Bernoulli}(p)$

Problema: hallar un intervalo de confianza de nivel $(1 - \alpha)$ para $\theta = p$.

Si n es suficientemente grande por teorema central del límite (T.C.L.) sabemos si $X \sim B(n, p) \Rightarrow \frac{X}{n} = \hat{p} \sim N\left(p, \frac{p \cdot q}{n}\right)$

Recordemos si $X \sim B(n, p)$

$$E(x) = n \cdot p$$

$$E\left(\frac{x}{n}\right) = \frac{1}{n} \cdot E(x) = \frac{1}{n} \cdot n \cdot p = p$$

$$V\left(\frac{x}{n}\right) = \frac{1}{n^2} \cdot V(x) = \frac{1}{n^2} \cdot n \cdot p \cdot q = \frac{p \cdot q}{n}$$

Luego $\frac{\hat{p} - p}{\underbrace{\sqrt{\frac{p \cdot q}{n}}}_{h(x_1, \dots, x_n, p)}} \sim N(0, 1)$

Fijado un nivel de confianza: $(1 - \alpha)$

$$\begin{aligned} |h(x_1, \dots, x_n, p)| &\leq z_{\frac{\alpha}{2}} \\ |\hat{p} - p|^2 &\leq \left(z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{p \cdot q}{n}} \right)^2 \\ \hat{p}^2 + p^2 - 2 \cdot \hat{p} \cdot p - z_{\frac{\alpha}{2}}^2 \cdot \frac{p \cdot q}{n} &\leq 0 \\ G(p) = p^2 \cdot \left[1 + z_{\frac{\alpha}{2}}^2 \cdot \frac{1}{n} \right] + p \cdot \left[-2 \cdot \hat{p} - z_{\frac{\alpha}{2}}^2 \cdot \frac{1}{n} \right] + \hat{p}^2 &\leq 0 \end{aligned}$$

$\forall p \in [r_1, r_2] \leq G(p) \leq 0 \Leftrightarrow r_1, r_2$ son raíces de $G(p)$

Obtengo las raíces:

$$\frac{2\hat{p} + z_{\frac{\alpha}{2}}^2 \frac{1}{n} + 2 \cdot z_{\frac{\alpha}{2}} \sqrt{\frac{z_{\frac{\alpha}{2}}^2}{4 \cdot n^2}}}{2 \cdot \left[1 + z_{\frac{\alpha}{2}}^2 + \frac{1}{n} \right]}$$

Es el intervalo de confianza de nivel aproximado $(1 - \alpha)$ para $\theta = p$

Si n es grande $\frac{z_{\frac{\alpha}{2}}^2}{n}$ es despreciable

Entonces IC se reduce a : $\boxed{\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}}$ ⁸

Elegir el n para que el IC tenga nivel aproximado $(1 - \alpha)$ y una longitud de a lo sumo w

$$2 \cdot z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}} \leq w$$

Puedo acotar $\hat{p} \cdot \hat{q} = \hat{p} * (1 - \hat{p}) \leq \frac{1}{4} \Rightarrow z_{\frac{\alpha}{2}} \sqrt{\frac{1}{n}} \leq w$

$$\boxed{\therefore n \geq \left(\frac{z_{\frac{\alpha}{2}}}{w} \right)^2}$$

⁸Intervalo de confianza tradicional.

Ejemplo

Sea X_1, \dots, X_n una M.A. $\sim P(\lambda)$ con n suficientemente grande.

Obtener un IC de nivel aproximado $(1 - \alpha)$ para λ

Nota Si $X \sim P(\lambda) \Rightarrow E(x) = V(x) = \lambda$

Por T.C.L $\underbrace{\frac{\bar{X} - \lambda}{\sqrt{\frac{\lambda}{n}}}}_{h(x_1, \dots, x_n, \lambda)} \sim N(0, 1)$

Fijamos un nivel aproximado $(1 - \alpha)$

$$\begin{aligned}(1 - \alpha) &= P\left(|h(x_1, \dots, x_n, \lambda)| \leq z_{\frac{\alpha}{2}}\right) \\&= (\bar{X} - \lambda \cdot \sqrt{n})^2 \leq (z_{\frac{\alpha}{2}} \cdot \sqrt{\lambda})^2 \\&= n \cdot (\bar{X}^2 + \lambda^2 - 2 \cdot \lambda \cdot \bar{X}) \leq z_{\frac{\alpha}{2}}^2 \cdot \lambda \\&= n \cdot \bar{X}^2 + n \cdot \lambda^2 - 2 \cdot n \cdot \lambda \cdot \bar{X} - z_{\frac{\alpha}{2}}^2 \cdot \lambda \leq 0\end{aligned}$$

Obtenemos un polinomio cuadrático en λ

$$n \cdot \bar{x}^2 + n \cdot \lambda^2 - 2 \cdot n \cdot \lambda \cdot \bar{x} - z_{\frac{\alpha}{2}}^2 \cdot \lambda \leq 0 \quad \forall \lambda : \lambda \in [x_1, x_2]$$

Donde x_1, x_2 son las raíces del polinomio

Luego el intervalo aproximado de nivel $(1 - \alpha)$ para λ es:

$$IC_{\lambda} = \frac{z_{\frac{\alpha}{2}}^2 + 2 \cdot n \cdot \bar{X} \pm z_{\frac{\alpha}{2}} \cdot \sqrt{z_{\frac{\alpha}{2}}^2 + 4 \cdot n \cdot \bar{X}}}{2 \cdot n}$$

7.3. Distribución t de Student con n grados de libertad

Sea X_1, \dots, X_n una M.A. $\sim N(\mu, \sigma^2)$ con μ y σ^2 desconocidos

- Si $n \geq 40$ entonces vimos que un IC de nivel de confianza aproximado $(1 - \alpha)$ para μ es: $\bar{X} \pm z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$

- ¿Qué pasa si $n < 40$?

Podemos afirmar que $\left(\frac{\bar{X} - \mu}{S/\sqrt{n}}\right) \not\sim N(0, 1)$

Entonces necesitamos otro estimador para obtener un IC para μ en esta situación.

Definición: sean $Z \sim N(0, 1)$ e $Y \sim \chi^2_\nu$: con Z e Y independientes.

La variable aleatoria $\frac{Z}{\sqrt{\frac{Y}{\nu}}} = T$ tiene una distribución **t-student** con ν grados de libertad .

Notación: $T \sim t_\nu$

7.3.1. Características de la t de Student

- Es simétrica en torno del origen y tiene forma de campana.
- Si ν es pequeño entonces la t_ν es más dispersa que una $N(0, 1)$
- $P(T \leq t) \xrightarrow{\nu \rightarrow \infty} \Phi(t)$

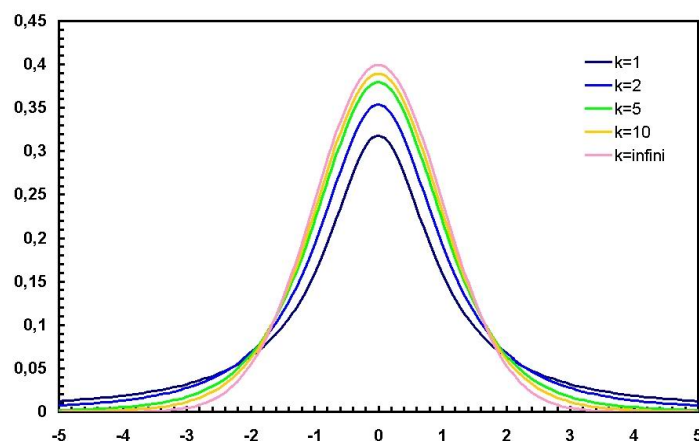


Figura 6: Gráfico de la t -student con k grados de libertad

Teorema Sea X_1, \dots, X_n una M.A. $\sim N(\mu, \sigma^2)$ entonces:

1. $\left(\frac{\bar{X}-\mu}{\sigma}\right) \cdot \sqrt{n} = Z \sim N(0, 1)$
2. $\frac{(n-1) \cdot S^2}{\sigma^2} = Y \sim \chi_{n-1}^2$
3. Z e Y independientes entonces :

$$\frac{\left(\frac{\bar{X}-\mu}{\sigma}\right) \cdot \sqrt{n}}{\sqrt{\frac{S^2}{\sigma^2}}} = \frac{(\bar{X}-\mu) \cdot \sqrt{n}}{S} \sim t_{n-1}$$

Luego un IC de nivel de confianza $(1 - \alpha)$ para μ es $\boxed{\bar{X} \pm t_{\frac{\alpha}{2}, n-1} \cdot \frac{S}{\sqrt{n}}}$

Resumen

Sea X_1, \dots, X_n una M.A. $\sim N(\mu, \sigma^2)$ entonces:

- Si σ es conocido $\Rightarrow IC = \bar{X} \pm z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$
- Si σ es desconocido:
 - si $n \geq 40 \Rightarrow IC = \bar{X} \pm z_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$
 - si $n < 40 \Rightarrow IC = \bar{X} \pm t_{\frac{\alpha}{2}, (n-1)} \cdot \frac{S}{\sqrt{n}}$

Sea X_1, \dots, X_n una M.A. con $E(x) = \mu$ y $V(x) = S^2$

- Si n es suficientemente grande: $IC = \bar{X} \pm z_{\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$
- Si $X_i \sim Bernoulli(p)$ entonces IC para p es: $IC = \hat{p} \pm z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}$

7.4. Intervalo de confianza para la varianza

Sea X_1, \dots, X_n una M.A. $\sim N(\mu, \sigma^2)$ ⁹ entonces: $\underbrace{\frac{(n-1) \cdot S^2}{\sigma^2}}_{h(x_1, \dots, x_n, \sigma^2)} \sim \chi_{n-1}^2$

Fijado un nivel de confianza $(1 - \alpha)$ quiero obtener $a < b$ tales que:

$$\begin{aligned} 1 - \alpha &= P(a \leq h(x_1, \dots, x_n, \sigma^2) \leq b) \\ &= \chi_{(1-\frac{\alpha}{2}), n-1}^2 \leq \frac{(n-1) \cdot S^2}{\sigma^2} \leq \chi_{\frac{\alpha}{2}, n-1}^2 \\ &= \frac{(n-1) \cdot S^2}{\chi_{\frac{\alpha}{2}, n-1}^2} \leq \sigma^2 \leq \frac{(n-1) \cdot S^2}{\chi_{(1-\frac{\alpha}{2}), n-1}^2} \end{aligned}$$

\therefore Un IC de nivel $(1 - \alpha)$ para σ^2 es : $\left[\frac{(n-1) \cdot S^2}{\chi_{\frac{\alpha}{2}, n-1}^2}, \frac{(n-1) \cdot S^2}{\chi_{(1-\frac{\alpha}{2}), n-1}^2} \right]$

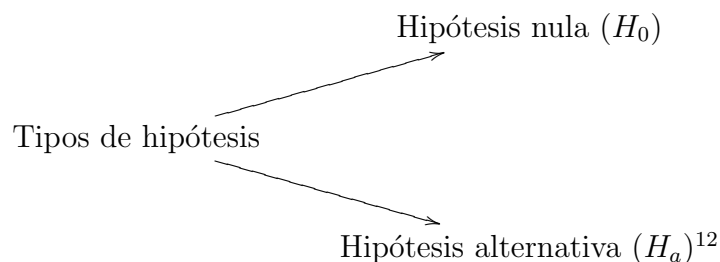
Y para σ es: $\left[\sqrt{\frac{(n-1) \cdot S^2}{\chi_{\frac{\alpha}{2}, n-1}^2}}, \sqrt{\frac{(n-1) \cdot S^2}{\chi_{(1-\frac{\alpha}{2}), n-1}^2}} \right]$

⁹Importante debe ser Normal la distribución sino no podemos saber nada.

8. Prueba de hipótesis basada en una muestra

Llamaremos con θ al parámetro poblacional.

Las hipótesis que el investigador debe establecer son dos:



Estas dos hipótesis deben ser contrapuestas.

8.1. Componentes de una prueba de hipótesis

8.1.1. Tipos de hipótesis alternativas (H_a)

- $H_a : \theta < \theta_0$ Unilateral de cola inferior.
- $H_a : \theta > \theta_0$ Unilateral de cola superior.
- $H_a : \theta \neq \theta_0$ Bilateral o de dos colas.

Con θ_0 valor fijado por el investigador.

El otro elemento que forma parte de una prueba de hipótesis es el **Estadístico de prueba** que es una función de la muestra aleatoria con distribución conocida.

8.1.2. Región de rechazo de la prueba (RR)

Son todos los valores del estadístico de prueba para el cual se rechaza la hipótesis nula.

Luego la H_0 será rechazada si el valor observado del estadístico de prueba está en la región de rechazo.

8.1.3. Tipos de errores de una prueba de hipótesis

Error de tipo I: (α) consiste en rechazar H_0 cuando H_0 es verdadera.

Error de tipo II: (β) consiste en aceptar H_0 cuando H_0 es falsa.

8.1.4. Nivel de significancia

Definimos el nivel de significancia de una prueba de hipótesis al supremo $P_\theta(\text{error de tipo I}) \forall \theta \in H_0$

8.2. Pasos a seguir para hacer una prueba de hipótesis

1. Identificar el parámetro de interés (θ).
2. Plantear las hipótesis nulas y alternativas pertinentes.
3. Indicar el estadístico de prueba.
4. Obtener la región de rechazo para un nivel de significancia α : (RR_α).
5. Evaluar el estadístico de prueba con los datos observados.
6. Tomar una decisión.
 - Si el valor observado está en la RR_α diremos que se rechaza H_0 con un nivel de significancia α .
 - En caso contrario diremos que no hay evidencia suficiente para rechazar H_0 a un nivel de significancia α .

8.3. Prueba de hipótesis para la media poblacional

- Caso A: sea X_1, \dots, X_n una M.A. $\sim N(\mu, \sigma^2)$ con σ conocido

$$H_0 : \mu = \mu_0$$

$$H_a : \mu \neq \mu_0$$

Estadístico de prueba: $\left(\frac{\bar{X} - \mu_0}{\sigma}\right) \cdot \sqrt{n} \sim N(0, 1)$ bajo H_0 .

Un criterio razonable para rechazar H_0 será si $\bar{x} \leq K_1^\alpha$ ó $\bar{x} \geq K_2^\alpha$

Fijado un nivel de significancia $\alpha \Rightarrow K_1^\alpha \stackrel{?}{=}$

$$\alpha = P(\text{error de tipo I}) = P(\bar{x} \leq K_1^\alpha \cup \bar{x} \geq K_2^\alpha) \text{ cuando } \mu = \mu_0$$

$$\alpha = P(\bar{x} \leq K_1^\alpha \text{ con } \mu = \mu_0) + P(\bar{x} \geq K_2^\alpha \text{ con } \mu = \mu_0)$$

$$\alpha = P\left(Z \leq \underbrace{\left(\frac{K_1^\alpha - \mu_0}{\sigma}\right) \cdot \sqrt{n}}_{z_1}\right) + P\left(Z \geq \underbrace{\left(\frac{K_2^\alpha - \mu_0}{\sigma}\right) \cdot \sqrt{n}}_{z_2}\right)$$

$$\begin{aligned}
RR_\alpha &= \{\bar{X} \leq -z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} + \mu_0 \text{ ó } \bar{X} \geq z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} + \mu_0\} \\
&= \{z \leq -z_{\frac{\alpha}{2}} \text{ ó } z \geq z_{\frac{\alpha}{2}}\} \\
&= \{|z| \geq z_{\frac{\alpha}{2}}\}
\end{aligned}$$

La función β es:

- $\beta(\mu) = \Phi(z_{\frac{\alpha}{2}} + (\frac{\mu_0 - \mu}{\sigma}) \cdot \sqrt{n}) - \Phi(-z_{\frac{\alpha}{2}} + (\frac{\mu_0 - \mu}{\sigma}) \cdot \sqrt{n}) \quad \forall \mu \neq \mu_0$
- β es simétrica en torno de μ id est: $\beta(\mu - h) = \beta(\mu + h) \quad \forall h \in \mathbb{R}$
- $\lim_{\mu \rightarrow \mu_0^-} \beta(\mu) = \lim_{\mu \rightarrow \mu_0^+} \beta(\mu) = 1 - \alpha$
- $\lim_{\mu \rightarrow \infty} \beta(\mu) = \lim_{\mu \rightarrow -\infty} \beta(\mu) = 0$
- β es creciente para $\mu < \mu_0$ y decreciente para $\mu > \mu_0$
- ¿Qué condición debe cumplir el n si $\mu' > \mu_0 \Rightarrow \beta(\mu') \leq \beta_0$?
 $\beta(\mu') = \Phi(z_{\frac{\alpha}{2}} + (\frac{\mu_0 - \mu'}{\sigma}) \cdot \sqrt{n}) - \Phi(-z_{\frac{\alpha}{2}} + (\frac{\mu_0 - \mu'}{\sigma}) \cdot \sqrt{n}) \quad \forall \mu' \neq \mu_0$
 $\beta(\mu') \leq \Phi(z_{\frac{\alpha}{2}} + (\frac{\mu_0 - \mu'}{\sigma}) \cdot \sqrt{n}) \leq \beta_0$

$$z_{\frac{\alpha}{2}} + (\frac{\mu_0 - \mu'}{\sigma}) \cdot \sqrt{n} \leq \Phi^{-1}(\beta_0)$$

$$z_{\frac{\alpha}{2}} - \frac{(-\mu_0 + \mu')}{\sigma} \cdot \sqrt{n} \leq -z_{\beta_0}$$

$$z_{\frac{\alpha}{2}} + z_{\beta_0} \leq \left(\frac{\mu' - \mu_0}{\sigma} \right) \cdot \sqrt{n}$$

$$\therefore n \geq \left[\frac{(z_{\frac{\alpha}{2}} + z_{\beta_0}) \cdot \sigma}{\mu' - \mu_0} \right]^2$$

Resumen

Sea X_1, \dots, X_n una M.A. $\sim N(\mu, \sigma^2)$ con σ conocido y $H_0 : \mu = \mu_0$

Estadístico de prueba: $\left(\frac{\bar{X} - \mu_0}{\sigma}\right) \cdot \sqrt{n} \sim N(0, 1)$ bajo H_0

Fijado un nivel α

H_a	RR_α	$\beta(\mu)$	$\beta(\mu') \leq \beta_0$
$\mu < \mu_0$	$\{z \leq -z_\alpha\}$	$1 - \Phi(-z_\alpha + (\frac{\mu_0 - \mu}{\sigma})\sqrt{n})$	$n \geq \left[\frac{\sigma \cdot (z_\alpha + z_{\beta_0})}{\mu' - \mu_0}\right]^2$
$\mu > \mu_0$	$\{z \geq z_\alpha\}$	$\Phi(z_\alpha + (\frac{\mu_0 - \mu}{\sigma})\sqrt{n})$	$n \geq \left[\frac{\sigma \cdot (z_\alpha + z_{\beta_0})}{\mu' - \mu_0}\right]^2$
$\mu \neq \mu_0$	$\{ z \geq z_{\frac{\alpha}{2}}\}$	$\Phi(z_{\frac{\alpha}{2}} + (\frac{\mu_0 - \mu}{\sigma})\sqrt{n}) - \Phi(-z_{\frac{\alpha}{2}} + (\frac{\mu_0 - \mu}{\sigma})\sqrt{n})$	$n \geq \left[\frac{\sigma \cdot (z_{\frac{\alpha}{2}} + z_{\beta_0})}{\mu' - \mu_0}\right]^2$

- Caso b: Sea X_1, \dots, X_n una M.A. con $E(x) = \mu$ y $V(x) = \sigma^2$ y n suficientemente grande.

Por T.C.L. $\bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \Leftrightarrow \left(\frac{\bar{X} - \mu}{\sigma}\right) \cdot \sqrt{n} \sim N(0, 1)$

- $\boxed{n \geq 30}$ y σ conocido entonces el estadístico de prueba es:

$$\left(\frac{\bar{X} - \mu_0}{\sigma}\right) \cdot \sqrt{n} \sim N(0, 1) \text{ bajo } H_0$$

- $\boxed{n \geq 40}$ y σ desconocido entonces el estadístico de prueba es:

$$\left(\frac{\bar{X} - \mu_0}{S}\right) \cdot \sqrt{n} \sim N(0, 1) \text{ bajo } H_0$$

- Las RR_α son las misma que en el **caso a**.

- Caso c: Sea X_1, \dots, X_n una M.A. $\sim N(\mu, \sigma^2)$ con σ desconocido y $H_0 : \mu = \mu_0$

- $\boxed{n \geq 40}$ Entonces el estadístico de prueba es:

$$\left(\frac{\bar{X} - \mu_0}{S}\right) \cdot \sqrt{n} \sim N(0, 1) \text{ bajo } H_0$$

Luego la región de rechazo para la prueba es la misma que la del **caso a**.

- $n < 40$ Entonces el estadístico de prueba va a ser :

$$\left(\frac{\bar{X} - \mu_0}{S} \right) \cdot \sqrt{n} \sim t_{n-1} \text{ bajo } H_0$$

Fijado un nivel de significancia α

$$RR_\alpha = \{\bar{x} \geq K_\alpha\} = \left\{ \left(\frac{\bar{x} - \mu_0}{S} \right) \cdot \sqrt{n} = t \geq \left(\frac{K_\alpha - \mu_0}{S} \right) \cdot \sqrt{n} \right\}$$

$$\alpha = P(\bar{x} \geq K_\alpha \text{ cuando } \mu = \mu_0)$$

$$\alpha = P \underbrace{\left(\frac{\bar{x} - \mu_0}{S} \right) \cdot \sqrt{n} \geq}_{t_{n-1}} \underbrace{\left(\frac{K_\alpha - \mu_0}{S} \right) \cdot \sqrt{n}}_{t^*}$$

$$t^* = t_{\alpha, n-1}$$

$$\alpha = t \geq t_{\alpha, n-1}$$

H_a	RR_α
$\mu < \mu_0$	$\{t \leq -t_{\alpha, n-1}\}$
$\mu > \mu_0$	$\{t \geq t_{\alpha, n-1}\}$
$\mu \neq \mu_0$	$\{ t \geq t_{\frac{\alpha}{2}, n-1}\}$

Ejemplo

Dos empresas distintas desean establecerse en cierta región y brindar servicios de televisión por cable. Se denota por p la proporción de suscriptores potenciales que prefieren la primera empresa sobre la segunda.

Se toma una muestra aleatoria de tamaño 25 para probar: $H_0 : p = 0,5$ contra $H_a : p \neq 0,5$.

Se representa con X el número de suscriptores en la muestra que está a favor de la primera empresa y con x el valor observado de X .

1. ¿Cuál de las siguientes regiones de rechazo es la más adecuada?

- $R_1 = \{x : x \leq 7 \text{ ó } x \geq 18\}$
- $R_2 = \{x : x < 8\}$
- $R_3 = \{x : x > 17\}$

La más adecuada es R_1 que corresponde a una H_a de dos colas.

2. ¿Cuál es la distribución de probabilidad del estadístico de prueba X cuando H_0 es verdadera? Utilícela para calcular la probabilidad de error de tipo I

Mi estadístico de prueba va a ser $X \sim B(25; 0,5)$ bajo H_0

$$\begin{aligned}
 P(\text{error de tipo I}) &= P(x \leq 7 \cup x \geq 18 \text{ cuando } p = p_0) \\
 &= P(x \leq 7 \text{ cuando } p = p_0) + P(x \geq 18 \text{ cuando } p = p_0) \\
 &= B(7; 25 : 0,5) + 1 - P(x \leq 17 \text{ cuando } p = p_0) \\
 &= B(7; 25 : 0,5) + 1 - B(17; 25; 0,5)
 \end{aligned}$$

$$\begin{aligned}
 P(\text{error de tipo I}) &= P(x \leq 7) + 1 - P(x \leq 17) \text{ cuando } p = p_0 \\
 &= \Phi\left(\frac{7 - n \cdot p_0}{\sqrt{n \cdot p_0 \cdot q_0}}\right) + 1 - \Phi\left(\frac{17 - n \cdot p_0}{\sqrt{n \cdot p_0 \cdot q_0}}\right) \\
 &= \Phi(-2,2) + 1 - \Phi(1,8) \\
 &= 0,0139 + 10,9641 \\
 &= 0,0498
 \end{aligned}$$

3. Calcule la probabilidad de error de tipo II para la región seleccionada cuando $p = 0,3$.

Repita lo mismo para $p = 0,4$; $p = 0,6$ y $p = 0,7$.

$$\beta(p) = P(7 < x \leq 17) = \sum_{k=8}^{17} \binom{n}{k} \cdot p^k \cdot q^{n-k}$$

p	$\beta(p)$	$\beta(17,25,p) - \beta(7,25,p)$
0,3	0,488	
0,4	0,845	
0,6	0,845	
0,7	0,488	

4. Mediante la región seleccionada ¿qué concluye si 6 de los 25 individuos favorecieron a la primera empresa?

8.4. Prueba de hipótesis para la varianza poblacional

Sea X_1, \dots, X_n M.A. $\sim N(\mu, \sigma^2)$ con σ desconocido y $H_0 : \sigma^2 = \sigma_0^2$.

El estimador de prueba va a ser: $\frac{(n-1).S^2}{\sigma^2} \sim \chi_{n-1}^2$ bajo H_0

Un criterio razonable para rechazar H_0 en favor de la hipótesis alternativa será:

$$S^2 \geq K_\alpha \Leftrightarrow \frac{(n-1).S^2}{\sigma_0^2} \geq \frac{K_\alpha.(n-1)}{\sigma_0^2}$$

$$\alpha = P(S^2 \geq K_\alpha \text{ cuando } \sigma^2 = \sigma_0^2)$$

$$\alpha = P\left(\frac{(n-1).S^2}{\sigma_0^2} \geq \frac{(n-1).K_\alpha}{\sigma_0^2}\right)$$

$$RR_\alpha = \{\chi^2 \geq \chi_{\alpha, n-1}^2\}$$

En síntesis

H_a	RR_α
$\sigma^2 > \sigma_0^2$	$\{\chi^2 \geq \chi_{\alpha, n-1}^2\}$
$\sigma^2 < \sigma_0^2$	$\{\chi^2 \leq \chi_{1-\alpha, n-1}^2\}$
$\sigma^2 \neq \sigma_0^2$	$\{\chi^2 \geq \chi_{\frac{\alpha}{2}, n-1}^2 \text{ ó } \chi^2 \leq \chi_{1-\frac{\alpha}{2}, n-1}^2\}$

8.5. Prueba de hipótesis para la proporción poblacional

Sea X_1, \dots, X_n una M.A. $\sim \text{Bernoulli}(p)$ con $H_0 : p = p_0$

Con p_0 valor fijado por el investigador.

Sabemos que: $\sum_{i=1}^n x_i \sim B(n, p)$

- Caso a: n pequeño, el estadístico de prueba va a ser:

$$X = \sum_{i=1}^n x_i \sim B(n, p_0)$$

H_a	RR_α
$p > p_0$	$\{x : x \geq K_\alpha\}$
$p < p_0$	$\{x : x \leq K_\alpha\}$
$p \neq p_0$	$\{x : x \leq K_1^\alpha \text{ ó } x \geq K_2^\alpha\}$

Sea $x \in \mathbb{Z}^{\geq 0}$

Definimos $B(x; n, p) = \sum_{k=0}^x \binom{n}{k} \cdot p^k (1-p)^{n-k}$

- En el caso $H_a : p > p_0$ y supongamos $K_\alpha \in \mathbb{Z}^{\geq 0}$

$$\begin{aligned} P(\text{error de tipo I}) &= P(X \geq K_\alpha \text{ cuando } p = P_0) \\ &= 1 - P(X \leq K_{\alpha-1} \text{ cuando } p = p_0) \\ &= 1 - B(K_{\alpha-1}; n, p_0) \leq \alpha \end{aligned}$$

- En el caso $H_a : p < p_0$

$$\begin{aligned} P(\text{error de tipo I}) &= P(X \leq K_\alpha \text{ cuando } p = P_0) \\ &= B(K_\alpha; n, p_0) \leq \alpha \end{aligned}$$

- En el caso $H_a : p \neq p_0$

$$P(\text{error de tipo I}) = P(X \leq K_1^\alpha : p = P_0) + P(X \geq K_2^\alpha : p = P_0)$$

- Caso b: cuando n es suficientemente grande:

$$X \sim N(n.p, n.p.q) \xrightarrow{T.C.L} \frac{X - n.p}{\sqrt{n.p.q}} \sim N(0, 1)$$

El estadístico de prueba va a ser: $\frac{X - n.p_0}{\sqrt{n.p_0.q_0}} \sim N(0, 1)$ Bajo H_0

Luego fijado el nivel de significancia α

H_a	RR_α	$\beta(p)$
$p > p_0$	$\{z \geq z_\alpha\}$	$\Phi(z_\alpha \sqrt{\frac{p_0 \cdot q_0}{p \cdot q}} + \sqrt{n} \cdot \frac{p_0 - p}{\sqrt{p \cdot q}})$
$p < p_0$	$\{z \leq -z_\alpha\}$	$1 - \Phi(-z_\alpha \sqrt{\frac{p_0 \cdot q_0}{p \cdot q}} + \sqrt{n} \cdot \frac{p_0 - p}{\sqrt{p \cdot q}})$
$p \neq p_0$	$\{ z \geq z_{\frac{\alpha}{2}}\}$	$\Phi(z_{\frac{\alpha}{2}} \sqrt{\frac{p_0 \cdot q_0}{p \cdot q}} + \sqrt{n} \cdot \frac{p_0 - p}{\sqrt{p \cdot q}}) + \Phi(-z_{\frac{\alpha}{2}} \sqrt{\frac{p_0 \cdot q_0}{p \cdot q}} + \sqrt{n} \cdot \frac{p_0 - p}{\sqrt{p \cdot q}})$

Ejemplo

(Extraído de libro Probabilidad y estadística de Jay L. Devore).

Muchos consumidores están volviendo a los medicamentos genéricos como una forma de reducir el costo de las medicaciones prescritas.

Se proporcionan los resultados de 102 médicos, de los cuales sólo 47 de los encuestados conocían el nombre genérico para el fármaco metadona.

¿Proporciona lo anterior evidencia suficiente para concluir que menos de la mitad de los médicos conocen el nombre genérico para la metadona?

Lleve a cabo una prueba de hipótesis con un nivel de significancia 0,05.
 $\theta = p \rightarrow$ proporción de médicos que conocen el nombre genérico de la metadona.

$$H_a : p < 0,5 = P_0$$

$$\alpha = 0,05$$

$$P(\text{error de tipo I}) = B(K_\alpha^*; 102; 0,5)$$

Caso B: Proporción

como n es suficientemente grande mi estadístico de prueba va a ser:

$$\frac{X - n \cdot p_0}{\sqrt{n \cdot p_0 \cdot q_0}} \sim N(0, 1) \text{ Bajo } H_0$$

$$\alpha = 0,05 \Rightarrow RR_{0,05} = \{z : z < -0,05\} = \frac{1,64+1,65}{2} = 0,645$$

$$z_{obs} = \frac{47 - 102 * 0,5}{\sqrt{102 * 0,5 * 0,5}} = -0,79 \notin RR_{0,05}$$

\therefore No hay evidencia suficiente para rechazar H_0 al 5 %

8.6. P-Valor

Definición: llamaremos p valor o nivel de significancia alcanzado al menor nivel de significancia a partir del cual rechazamos H_0 con los resultados obtenidos en la muestra.

- $p \text{ valor} \leq \alpha \Rightarrow$ rechazo a un nivel α

$$\text{Id est: } z_\alpha \leq z_{obs}$$

- $p \text{ valor} > \alpha \Rightarrow$ no hay evidencia suficiente para rechazar H_0

$$\text{Id est: } z_{obs} \notin RR_\alpha$$

Estadístico de prueba	H_a	p valor
Z	cola superior	$P(Z \geq z_{obs})$
	cola inferior	$P(Z \leq z_{obs})$
	bilateral	$2 \cdot P(Z \geq z_{obs})$
T	cola superior	$P(T \geq t_{obs})$
	cola inferior	$P(T \leq t_{obs})$
	bilateral	$2 \cdot P(T \geq t_{obs})$

9. Inferencia basada en dos muestras

No lo vimos ☺