Tecnológico de Monterrey

Concentración en Analítica de Datos y Herramientas de Inteligencia Artificial II

Daniel Otero Fadul

Departamento de Ciencias
Escuela de Ingeniería y Ciencias

Axiomas de la Probabilidad

Un **experimento aleatorio** es un procedimiento bien definido que produce un resultado observable que no se puede predecir con antelación. Sea Ω el conjunto que contiene todos los resultados posibles de un experimento aleatorio; a este conjunto se le conoce como **espacio muestral**. Cualquier subconjunto de Ω se le llama **evento**.

Axiomas de la Probabilidad

Por ejemplo, lanzar una moneda dos veces es un experimento aleatorio que tiene el siguiente espacio muestral:

$$\Omega = \{SS, SC, CS, CC\},$$

donde S y C hacen referencia a que un lanzamiento haya sido "sello" o "cara", respectivamente. Un posible evento de este experimento es $A = \{CC\}$.

Nótese que, ya que los eventos son conjuntos, todas las operaciones de conjuntos son válidas también para eventos.

Axiomas de la Probabilidad

La probabilidad de un evento A es una medida de qué tan probable es que ocurra dicho evento. Esta medida debe cumplir con los siguientes axiomas:

- $\forall A \subseteq \Omega,\ P(A) \in [0,1]$.
- $P(\Omega) = 1$.
- Para toda colección contable de eventos disyuntos $\{A_i\}_{i=1}^{n}$, $n \in \mathbb{N}$, $A_i \cap A_j = \emptyset$ $\forall i \neq j$,

$$P\left(\cup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$$

Axiomas de la Probabilidad

Estos axiomas se utilizan para definir y probar distintos resultados en probabilidad. Por ejemplo, podemos calcular la probabilidad del complemento de un evento A:

$$\begin{aligned} 1 &= P(\Omega) \\ &= P(A \cup A^c) \\ &= P(A) + P(A^c). \end{aligned}$$

Por lo tanto, $P(A^c) = 1 - P(A)$.

Sea $A = \{CC\}$. Asumiendo que cada posible resultado de Ω tiene la misma probabilidad de ocurrir, tenemos que

$$\begin{aligned} P(A^c) &= 1 - P(A) \\ &= 1 - \frac{1}{4} = \frac{3}{4}. \end{aligned}$$

Probabilidad Condicional

Sean A y B dos eventos que pertenecen a un espacio muestral Ω. La probabilidad de que ocurra el vento A dado que se sabe que el evento B ocurrió se define como

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Si A y B son independientes, se tiene que $P(A|B) = P(A)$, lo que implica que $P(A \cap B) = P(A)P(B)$.

Probabilidad Condicional

Retomemos el experimento de la moneda. El espacio muestral está dado por $\Omega = \{SS, SC, CS, CC\}$. Sea A el evento de que al menos un lanzamiento fue cara, y sea B el evento de que al menos uno de dos lanzamientos fue sello. ¿A qué es igual $P(A|B)$?

Tenemos que $A = \{SC, CS, CC\}$ y $B = \{SS, SC, CS\}$. Por lo tanto, $A \cap B = \{SC, CS\}$. Entonces,

$$\begin{aligned} P(A|B) &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{P(\{SC\}) + P(\{CS\})}{P(\{SS\}) + P(\{SC\}) + P(\{CS\})} \\ &= \frac{\frac{1}{2}}{\frac{3}{4}} = \frac{2}{3}. \end{aligned}$$

Probabilidad Total

Sea $\{A_i\}_{i=1}^{n}$ una colección de eventos mutuamente excluyentes que forman una partición del espacio muestral: $A_i \cap A_j = \emptyset$, $\forall i \neq j$ y $\cup_{i=1}^{n} A_i = \Omega$. Entonces, para cualquier evento $B \subset \Omega$, tenemos que

$$\begin{aligned} P(B) &= P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_n \cap B) \\ &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \cdots + P(B|A_n)P(A_n). \end{aligned}$$

Nótese que $\forall i,\ P(A_i) > 0$.

Probabilidad Total

Papá Noel empaca los juguetes que reparte en Navidad en cajas que contienen veinte juguetes. Suponga que el 60% de todas las cajas que reparte Papá Noel no contienen juguetes defectuosos, 30% contienen sólo un juguete defectuoso, y 10% contienen dos juguetes defectuosos. Si se elige una caja al azar, ¿cuál es la probabilidad de elegir aleatoriamente, sin reemplazo, dos juguetes sin defectos?

Probabilidad Total

Sean C_0, C_1 y C_2 los eventos de caja con cero juguetes defectuosos, caja con un juguete defectuoso y caja con dos juguetes defectuosos, respectivamente. Sea J_0 el evento de sacar dos juguetes sin defectos de una caja. Entonces,

$$
\begin{aligned}
P(J_0) &= P(J_0|C_0)P(C_0) + P(J_0|C_1)P(C_1) + P(J_0|C_2)P(C_2) \\
&= (1)\left(\frac{6}{10}\right) + \left(\frac{19}{20}\frac{18}{19}\right)\left(\frac{3}{10}\right) + \left(\frac{18}{20}\frac{17}{19}\right)\left(\frac{1}{10}\right) \\
&= \frac{903}{950} = 0.9505\ldots
\end{aligned}
$$

Teorema de Bayes

Sea $\{A_i\}_{i=1}^n$ una colección de eventos mutuamente excluyentes que forman una partición del espacio muestral: $A_i \cap A_j = \emptyset$, $\forall i \neq j$, $\cup_{i=1}^n A_i = \Omega$ y $P(A_i) > 0 \ \forall i$. Entonces, para cualquier evento $B \subset \Omega$ tal que $P(B) > 0$, tenemos que

$$
\begin{aligned}
P(A_i|B) &= \frac{P(B|A_i)P(A_i)}{P(B)} \\
&= \frac{P(B|A_i)P(A_i)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \cdots + P(B|A_n)P(A_n)}.
\end{aligned}
$$

Teorema de Bayes

Cuando la partición de Ω es de solo dos elementos ($A \cup \bar{A} = \Omega$), tenemos el siguiente caso especial:

$$
\begin{aligned}
P(A|B) &= \frac{P(B|A)P(A)}{P(B)} \\
&= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})}.
\end{aligned}
$$

Teorema de Bayes

Sean C_0, C_1 y C_2 los eventos de caja con cero juguetes defectuosos, un juguete defectuoso y dos juguetes defectuosos, respectivamente. Sea J_0 el evento de sacar dos juguetes sin defectos de una caja. Entonces, para responder la primera pregunta, necesitamos calcular $P(C_0|J_0)$:

$$
\begin{aligned}
P(C_0|J_0) &= \frac{P(J_0|C_0)P(C_0)}{P(J_0|C_0)P(C_0)+P(J_0|C_1)P(C_1)+P(J_0|C_2)P(C_2)} \\
&= \frac{(1)\left(\frac{6}{10}\right)}{(1)\left(\frac{6}{10}\right)+\left(\frac{19}{20}\frac{18}{19}\right)\left(\frac{3}{10}\right)+\left(\frac{18}{20}\frac{17}{19}\right)\left(\frac{1}{10}\right)} \\
&= 0.631229\ldots
\end{aligned}
$$

Variables Aleatorias

Una variable aleatoria es una variable cuyos valores dependen de un fenómeno aleatorio. Más formalmente, una variable aleatoria $X : \Omega \to E$, donde E es algún espacio medible, es una función que le asigna a cada elemento del espacio muestral Ω un número entero o real. Por ejemplo, para el experimento de la moneda, podríamos definir a X como la siguiente variable aleatoria:

X
$SS \to 0$
$SC \to 1$
$CS \to 2$
$CC \to 3$

Variables Aleatorias Discretas

Una variable aleatoria es discreta si el conjunto de valores que toma es *contable*. Justamente, la variable que definimos en la diapositiva anterior es un ejemplo de variable aleatoria discreta.

Variables Aleatorias Discretas

Toda variable aleatoria discreta tiene asociada una **función de masa de probabilidad**, la cual le asigna probabilidades a cada valor que puede tomar la variable aleatoria. Esta función $f : E \to \mathbb{R}$ debe cumplir con las siguientes condiciones:

- $\forall x \in E,\ f(x) \geq 0$.
- $\sum_x f(x) = 1$.
- $P(X = x) = f(x)$.

Además, asociada a la función de masa de probabilidad tenemos la **función de distribución acumulada**:

$$F(x) = P(X \leq x) = \sum_{y \leq x} f(y).$$

Variables Aleatorias Discretas

Por ejemplo, volviendo a la moneda que lanzamos dos veces, la función de masa de probabilidad sería igual a

$$f(x) = \frac{1}{4}, \; x = 0, 1, 2, 3. \tag{1}$$

Nótese que $f(1) = P(X = 1) = P(\{SC\})$.

Variables Aleatorias Continuas

Una variable aleatoria es continua si el conjunto de valores que toma es *no contable*. Por ejemplo, la temperatura de una habitación o el tiempo que esperamos para la llegada de un bus en una estación son ejemplos de variables aleatorias continuas.

Variables Aleatorias Continuas

Toda variable aleatoria continua tiene asociada una **función de densidad de probabilidad**. Esta función $f : E \to \mathbb{R}$ debe cumplir con las siguientes condiciones:

- $\forall x \in E,\ f(x) \geq 0$.
- $\int_{\mathbb{R}} f(x)dx = 1$.
- $P(a < X < b) = \int_a^b f(x)dx$.

La función de densidad de probabilidad también tiene asociada una **función de distribución acumulada**:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(y)dy.$$

Nótese que

$$\frac{d}{dx}F(x) = f(x).$$

Variables Aleatorias Continuas

Consideremos la siguiente función de densidad de probabilidad:

$$f(x) = \begin{cases} \frac{1}{3}x^2, & -1 < x < 2 \\ 0, & x \notin (-1, 2) \end{cases} \tag{2}$$

¿A qué es igual $P(0 < X < 1)$?

$$\begin{aligned} P(0 < X < 1) &= \int_0^1 \frac{1}{3}x^2 dx \\ &= \frac{1}{3}\int_0^1 x^2 dx \\ &= \frac{1}{3}\left(\frac{x^3}{3}\Big|_0^1\right) \\ &= \frac{1}{9}. \end{aligned}$$

Valor Esperado

El valor esperado de una variable aleatoria discreta se define como

$$E(X) = \sum_{x} x f(x).$$

Este se puede entender como un promedio o como un parámetro de localización: el valor en el que nuestra función de masa de probabilidad está "centrada" o ubicada.

Valor Esperado

Volviendo a recordar la función de masa de probabilidad del experimento aleatorio de la moneda, tenemos que el valor esperado de X es igual a

$$\begin{aligned} E(X) &= \sum_{x=0}^{3} x f(x) \\ &= (0)\frac{1}{4} + (1)\frac{1}{4} + (2)\frac{1}{4} + (3)\frac{1}{4} \\ &= \frac{3}{2}. \end{aligned}$$

Valor Esperado

En el caso de una variable aleatoria continua tenemos que

$$E(X) = \int_{\mathbb{R}} xf(x)dx.$$

Vale la pena mencionar que el valor esperado también se conoce como el *momento de primer orden* de la variable aleatoria X.

Valor Esperado

Para la función de densidad de probabilidad que mencionamos anteriormente tenemos que

$$\begin{aligned} E(X) &= \int_{-1}^{2} x\left(\frac{1}{3}x^2\right)dx \\ &= \frac{1}{3}\int_{-1}^{2} x^3 dx \\ &= \frac{1}{3}\left(\frac{x^4}{4}\Big|_{-1}^{2}\right)dx \\ &= \frac{5}{4}. \end{aligned}$$

Valor Esperado

Vale la pena mencionar que el valor esperado tiene las siguientes propiedades:

- $\forall a \in \mathbb{R}$, $E(aX) = aE(X)$.
- Sean X_1 y X_2 variables aleatorias. Entonces, $E(X_1 + X_2) = E(X_1) + E(X_2)$.
- Para una variable aleatoria discreta X, $E(g(X)) = \sum_x g(x)f(x)$.
- De manera similar, para una variable aleatoria continua X, tenemos que $E(g(X)) = \int_{\mathbb{R}} g(x)f(x)dx$.

Varianza

La varianza de una variable aleatoria X es una medida de dispersión de esta alrededor de su media o valor esperado. Esta se define como

$$\begin{aligned} \mathrm{Var}(X) &= E((X - E(X))^2) \\ &= E(X^2) - (E(X))^2. \end{aligned}$$

Esta medida de dispersión tiene estas propiedades:

- $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$, donde $a, b \in \mathbb{R}$.
- Si X_1 y X_2 son variables aleatorias independientes,

$$\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2).$$

Varianza

Para la variable aleatoria X con función de masa de probabilidad definida en (1), tenemos que la varianza es igual a

$$\begin{aligned} \mathrm{Var}(X) &= E(X^2) - (E(X))^2 \\ &= \sum_{x=0}^{3} x^2 f(x) - \left(\frac{3}{2}\right)^2 \\ &= (0^2)\frac{1}{4} + (1^2)\frac{1}{4} + (2^2)\frac{1}{4} + (3^2)\frac{1}{4} - \frac{9}{4} \\ &= \frac{5}{4}. \end{aligned}$$

Varianza

La varianza de la variable aleatoria continua X, cuya función de densidad de probabilidad está definida en (2), es igual a

$$
\begin{aligned}
\mathrm{Var}(X) &= E(X^2) - (E(X))^2 \\
&= \frac{1}{3}\int_{-1}^{2} x^4 dx - \left(\frac{5}{4}\right)^2 \\
&= \frac{1}{3}\left(\left.\frac{x^5}{5}\right|_{-1}^{2}\right) - \frac{25}{16} \\
&= \frac{11}{5} - \frac{25}{16} = \frac{51}{80}.
\end{aligned}
$$

Distribución de Bernoulli

Para la distribución de Bernoulli tenemos una variable aleatoria X que solo toma dos valores: X es igual a 1 con probabilidad p e igual a 0 con probabilidad $1 - p$. Su función de masa de probabilidad es igual a

$$f(x) = \begin{cases} 1 - p, & x = 0 \\ p, & x = 1 \end{cases}$$

Distribuciones de Probabilidad

Además, tenemos que

$$\begin{aligned} E(X) &= \sum_{x=0}^{1} x f(x) \\ &= (0)(1-p) + (1)p = p. \end{aligned}$$

En cuanto a la varianza, está es igual a

$$\begin{aligned} \mathrm{Var}(X) &= \sum_{x=0}^{1} x^2 f(x) - p^2 \\ &= (0^2)(1-p) + (1^2)p - p^2 = p(1-p). \end{aligned}$$

Distribución Normal

La distribución más importante en estadística es la distribución normal, también conocida como distribución Gaussiana en honor a Karl Friedrich Gauss. La función de densidad de probabilidad de una variable aleatoria normal X está dada por

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

donde $\mu = E(X)$ y $\sigma^2 = \text{Var}(X)$.

Distribuciones de Probabilidad

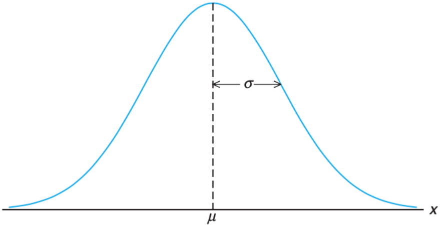

Figura: Gráfica de la distribución normal con parámetros μ y σ. A esta distribución se le denota algunas veces como $N(\mu, \sigma^2)$. Imagen tomada de [2].

Sea X una variable aleatoria normal con $\mu = 50$ y $\sigma = 10$. ¿Cuál es la probabilidad de que X tome valores entre 45 y 62?

$$
\begin{aligned}
P(45 < X < 62) &= P\left(\frac{45-50}{10} < \frac{X-\mu}{\sigma} < \frac{62-50}{10}\right) \\
&= P\left(-\frac{1}{2} < Z < \frac{6}{5}\right) \\
&= P\left(Z < \frac{6}{5}\right) - P\left(Z < -\frac{1}{2}\right) \\
&= 0.8849 - 0.3085 = 0.5764.
\end{aligned}
$$

Ley de los Grandes Números

Sea $\{X_i\}_{i=1}^{\infty}$ una colección de variables aleatorias independientes e idénticamente distribuidas. Si su momento de primer orden es finito, es decir, $\forall i \in \mathbb{N}$, $E(X_i) = \mu$, entonces

$$\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu$$

con probabilidad igual a uno a medida que $n \to \infty$. A este resultado se le conoce como la **ley de los grandes números**.

Teorema del Límite Central

Sea $\{X_i\}_{i=1}^{\infty}$ una colección de variables aleatorias independientes e idénticamente distribuidas. Si sus momentos de primer orden y sus varianzas son iguales y finitas, es decir, $\forall i \in \mathbb{N}$, $E(X_i) = \mu < \infty$ y $\text{Var}(X_i) = \sigma^2 < \infty$, entonces

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1),$$

donde $\bar{X}$ es igual a

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}.$$

Este resultado se le conoce como el **teorema del límite central**.

Teorema del Límite Central

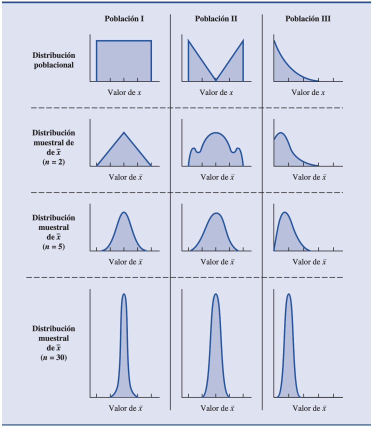

Figura: En este figura se ilustra cómo funciona el teorema del límite central en tres poblaciones diferentes. En la primera fila se puede ver que ninguna de las tres poblaciones está distribuida normalmente: los elementos de la población I tienen una distribución uniforme; los elementos de la segunda población II siguen una distribución que se le conoce como distribución en forma de orejas de conejo; la forma de la distribución de los elementos de la población III se asemeja a una distribución exponencial. Se puede ver que, independientemente de las distribuciones de los elementos de cada población, a medida que n aumenta la distribución de la muestra tiende a una distribución normal. Imagen tomada de [3].

Estimador Muestral de la Media Poblacional

Consideremos nuevamente el estimador muestral de la media de una población:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Ya que el valor de $\bar{x}$ depende de la muestra que tengamos, la cual se elige aleatoriamente, su valor va a ser diferente cada vez que la muestra cambie, razón por la cual el estimador se comporta como una variable aleatoria.

Estimador Muestral de la Media Poblacional

Dicho lo anterior, podemos ver que la ley de los grandes número nos dice que, si la población es infinita, el valor de $\bar{x}$ tiende al valor verdadero de la población a medida que el tamaño de la muestra n tiende a infinito:

$$\lim_{n\to\infty} \bar{x} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} x_i = \mu.$$

Estimador Muestral de la Media Poblacional

Es más, el teorema del límite central nos dice que

$$\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1).$$

Dado esto, ¿a qué serían iguales $E(\bar{x})$ y $\text{Var}(\bar{x})$?

Estimador Muestral de la Media Poblacional

Sea $Z \sim N(0,1)$. Entonces,

$$\begin{aligned}
\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} &= Z \\
\bar{x} &= \mu + \frac{\sigma}{\sqrt{n}} Z \\
E(\bar{x}) &= E\left(\mu + \frac{\sigma}{\sqrt{n}} Z\right) \\
E(\bar{x}) &= E(\mu) + E\left(\frac{\sigma}{\sqrt{n}} Z\right) \\
E(\bar{x}) &= \mu + \frac{\sigma}{\sqrt{n}} E(Z) \\
E(\bar{x}) &= \mu.
\end{aligned}$$

El hecho de que el valor esperado de $\bar{x}$ sea justamente igual al parámetro que se quiere estimar hace que $\bar{x}$ sea un **estimador insesgado** de la media poblacional μ. Dado esto, se dice que la media muestral es un estimador puntual de la media poblacional.

Estimador Muestral de la Media Poblacional

En cuanto a la varianza de $\bar{x}$ tenemos que

$$\begin{aligned} \text{Var}(\bar{x}) &= \text{Var}\left(\mu + \frac{\sigma}{\sqrt{n}} Z\right) \\ \text{Var}(\bar{x}) &= \left(\frac{\sigma}{\sqrt{n}}\right)^2 \text{Var}(Z) \\ \text{Var}(\bar{x}) &= \frac{\sigma^2}{n}. \end{aligned}$$

Nótese que a medida que n tiende a infinito la varianza tiende a cero. Esta característica de $\bar{x}$ lo convierte en un **estimador consistente**.

Estimador Muestral de la Media Poblacional

Vale la pena decir que, cuando la población consta de N elementos, la varianza de $\bar{x}$ es igual a

$$\text{Var}(\bar{x}) = \sqrt{\frac{N-n}{N-1}} \left(\frac{\sigma^2}{n} \right).$$

Al término que multiplica σ^2/n se le conoce como **factor de corrección para una población finita**. En múltiples ocasiones, cuando N es grande comparado con el tamaño de la muestra, este factor de correción toma un valor muy cercano a uno, razón por se considera que σ^2/n es una buena aproximación de la varianza de $\bar{x}$.

Estimador Muestral de la Media Poblacional

En el caso de poblaciones que siguen una distribución simétrica como la normal, los valores de la media y la mediana coinciden, sin embargo, como estimadores de este parámetro de la población, la desviación estándar de la mediana es mayor que la de la media, por lo cual se dice que la media muestral es más **eficiente** que la mediana muestral.

Estimador Muestral de la Varianza Poblacional

Como ya vimos, la varianza muestral es igual a

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

Esta expresión es un estimador puntual de la varianza poblacional.

Estimador Muestral de la Varianza Poblacional

Si las observaciones x_i provienen de una población normal, se puede demostrar que la expresión

$$\frac{n-1}{\sigma^2}s^2$$

sigue una distribución **chi-cuadrado** con $n-1$ grados de libertad.

Estimador Muestral de la Varianza Poblacional

La distribución chi-cuadrado con k grados de libertad se denota como χ^2_k. Si una variable aleatoria X posee esta distribución, su función de densidad de probabilidad está dada por

$$f(x) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}},\ x > 0,$$

donde Γ es la función gamma. En este caso, tenemos que $E(X) = k$ y $\text{Var}(X) = 2k$.

Estimador Muestral de la Varianza Poblacional

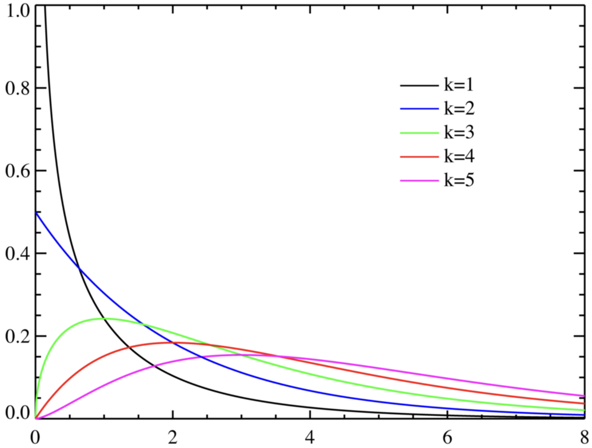

Figura: Gráfica de la distribución χ^2_k para distintos valores de k. Imagen tomada de `https://es.wikipedia.org/wiki/Distribuci%C3%B3n_%CF%87%C2%B2`.

Estimador Muestral de la Varianza Poblacional

Dado lo anterior, tenemos que

$$\begin{aligned} E\left(\frac{n-1}{\sigma^2}s^2\right) &= n-1 \\ \frac{n-1}{\sigma^2}E(s^2) &= n-1 \\ E(s^2) &= \sigma^2. \end{aligned}$$

Así que s^2 es un estimador insesgado de la varianza poblacional.

Estimador Muestral de la Varianza Poblacional

Por otro lado, en cuanto a la varianza de s^2, se tiene que

$$
\begin{aligned}
\text{Var}\left(\frac{n-1}{\sigma^2}s^2\right) &= 2(n-1) \\
\frac{(n-1)^2}{\sigma^4}\text{Var}(s^2) &= 2(n-1) \\
\text{Var}(s^2) &= 2\frac{\sigma^4}{n-1}.
\end{aligned}
$$

Por lo tanto, s^2 es un estimador consistente.

Estimación por Intervalo

Como se mencionó anteriormente, la media muestral $\bar{x}$ es un estimador puntual de la media poblacional μ. En general, como no se puede esperar que un estimador puntual suministre el valor exacto del parámetro poblacional, se suele calcular una **estimación por intervalo** al sumar y restar al estimador puntual una cantidad llamada **margen de error**. La fórmula general de una estimación por intervalo es

$$\text{estimación puntual} \;\pm\; \text{margen de error.}$$

Estimación por Intervalo

Por otro lado, queremos que en una estimación se tenga una idea de qué tan seguros podemos estar de que el intervalo que obtenemos contiene el verdadero valor del parámetro poblacional. A este grado de certeza se le llama **nivel de confianza** del intervalo, el cual posee un **coeficiente de confianza** que se define como $1 - \alpha$, donde $0 \leq \alpha \leq 1$.

Estimación por Intervalo

Un **intervalo de confianza del** $(\mathbf{1}-\alpha)\times\mathbf{100\%}$ tiene con un coeficiente de confianza igual a $1-\alpha$ y un nivel de confianza del $(1-\alpha)\times 100\%$.

Un típico valor de α es 0.05. En este caso decimos que tenemos un intervalo de confianza del 95% con un coeficiente de confianza igual a 0.95 y un nivel de confianza del 95%.

Media Poblacional: σ conocida

Para obtener un intervalo de confianza para la media poblacional se necesita la desviación estándar de la población ya que el margen de error depende de esta. Si la varianza de la población es conocida, tenemos que la estandarización del estimador de la media poblacional $\bar{x}$ sigue una distribución normal estándar:

$$\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1).$$

Media Poblacional: σ conocida

Teniendo en cuenta lo anterior, un intervalo con un coeficiente de confianza igual a $1 - \alpha$ será el rango de valores que concentre el $(1 - \alpha) \times 100\%$ de los valores que puede tomar la estandarización de la media muestral $\bar{x}$. Por lo tanto, el coeficiente de confianza se puede interpretar como una probabilidad, la cual es igual a

$$1 - \alpha = P\left(-z_{\alpha/2} \leq \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right), \tag{3}$$

donde $z_{\alpha/2}$ es el valor tal que

$$P\left(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \geq z_{\alpha/2}\right) = \frac{\alpha}{2}.$$

Media Poblacional: σ conocida

Dado lo anterior, la ecuación (3) se puede reescribir como

$$
\begin{aligned}
1-\alpha &= P\left(-z_{\alpha/2} \leq \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right) \\
&= P\left(-z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \bar{x}-\mu \leq z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(\bar{x}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right).
\end{aligned}
$$

Por lo tanto, el nivel de confianza $1-\alpha$ es equivalente a la probabilidad de que la media poblacional μ se encuentre en el intervalo

$$
\left[\bar{x}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right].
$$

Media Poblacional: σ conocida

Por lo tanto, tenemos que nuestro margen de error está dado por

$$\text{margen de error} = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}.$$

Así que los extremos de un intervalo de confianza de la media poblacional con un nivel de confianza del $(1-\alpha)\times 100\%$ son iguales a

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}.$$

Media Poblacional: σ conocida

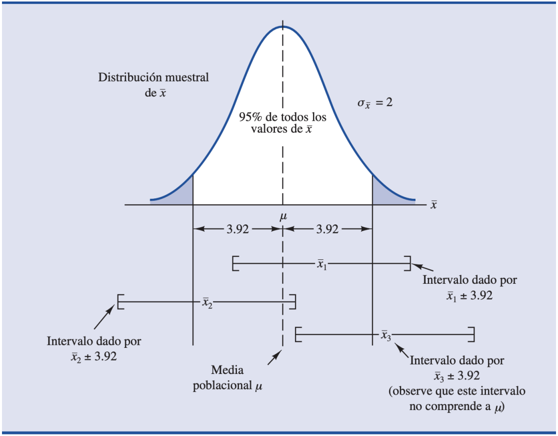

Figura: En esta figura se muestran tres intervalos de confianza del 95% obtenidos a partir de tres muestras diferentes. En este caso, $\sigma/\sqrt{n} = 2$ y $z_{\alpha/2} = 1.96$. Nótese que no siempre el intervalo de confianza contiene a la media poblacional μ, pero, en promedio, el 95% de los intervalos sí lo contendrá. Imagen tomada de [3].

Media Poblacional: σ conocida

Se obtuvo que la media muestral de una muestra de 40 artículos fue $\bar{x} = 25$. Si la desviación estándar de la población es $\sigma = 5$, ¿a qué es igual el intervalo de confianza de la media poblacional μ con $\alpha = 0.05$?

Media Poblacional: σ desconocida

Cuando la desviación estándar de la población no se conoce es necesario estimar su valor. En este caso, utilizamos el siguiente estimador de la desviación estándar poblacional

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

Media Poblacional: σ desconocida

Lo anterior implica que la estandarización de la media muestral está dada por

$$\frac{\bar{x} - \mu}{s/\sqrt{n}}.$$

Sin embargo este estadístico no sigue una distribución normal, sino una distribución **t de Student** de $n - 1$ grados de libertad.

Media Poblacional: σ desconocida

Sea X una variable aleatoria que sigue una distribución t de Student de k grados de libertad. Su función de densidad de probabilidad está dada por

$$f(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{k\pi}\,\Gamma\left(\frac{k}{2}\right)}\left(1+\frac{x^2}{k}\right)^{-\frac{k+1}{2}}, \quad -\infty < x < \infty.$$

Media Poblacional: σ desconocida

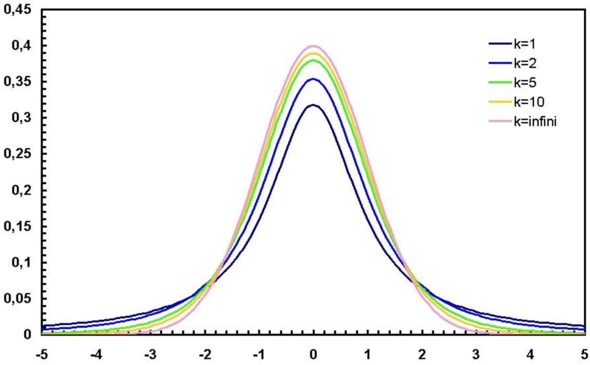

Figura: Gráfica de la distribución t de Student para distintos valores de k. Imagen tomada de `https://es.wikipedia.org/wiki/Distribuci%C3%B3n_t_de_Student`.

Media Poblacional: σ desconocida

En este caso tenemos que nuestro margen de error es igual a

$$\text{margen de error} = t_{\alpha/2}\frac{s}{\sqrt{n}}.$$

donde $t_{\alpha/2}$ es el valor tal que

$$P\left(\frac{\bar{x}-\mu}{s/\sqrt{n}} \geq t_{\alpha/2}\right) = \frac{\alpha}{2}.$$

Media Poblacional: σ desconocida

Así que los extremos de un intervalo de confianza de la media poblacional con un nivel de confianza del $(1-\alpha)\times 100\%$ son iguales a

$$\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}.$$

Media Poblacional: σ desconocida

Los datos muestrales siguientes provienen de una población normal: 10, 8, 12, 15, 13, 11, 6, 5. Para un nivel de confianza del 95%, ¿a qué es igual el intervalo de confianza de la media poblacional?

BIBLIOGRAFÍA

1 Ross S., *"Simulation"*, Quinta Edición, Elsevier, 2013.

2 Walpole R. E., Myers R. H., Myers S. L., Keying Y., *"Probability and Statistics for Engineers and Scientists"*, 9th Edition, Prentice Hall, 2012.

3 Anderson, D. R., Sweeney, D. J., Williams, T. A., *"Estadística para Administración y Economía"*, Décima Edición, CENGAGE Learning, 2008.