CAPÍTULO 6

DISTRIBUCIONES DE PROBABILIDAD CONTINUAS

6.1. Introducción

Como se dijo en el capitulo anterior, en muchas situaciones de la vida real surge la necesidad de dar respuesta a interrogantes como ¿Cuál es la probabilidad de un individuo en particular este entre 1,64 y 1,80 metros?, ¿Cuál es la probabilidad de que el ingreso promedio de los habitantes de determinada ciudad sea superior a 1500 bolívares fuertes?. Respuestas a estas interrogantes pueden ser dada haciendo uso de las distribuciones de probabilidad. En ambas situaciones se hace referencia a variables aleatorias continuas, por lo que para dar respuesta a dichas interrogantes se debe hacer uso de distribuciones de probabilidad continuas. Una distribución de probabilidad se llama continua si su función de distribución es continua.

En una distribución de probabilidad discreta un evento con probabilidad cero es un evento imposible. Esto no en el caso de una variable aleatoria continua, pues entre dos posibles valores de la misma hay infinitos valores posibles, aunque la probabilidad de ese intervalo no es cero, si lo es la de cada uno de esos valores individuales. Esto se debe al hecho de que la probabilidad de que X tome algún valor en un conjunto infinito como un intervalo, no puede calcularse mediante la adición simple de probabilidades de valores individuales. Realmente, cada valor tiene una probabilidad de ocurrencia infinitesimal, tan pequeña que estadísticamente equivale a cero.

Existen muchas leyes de probabilidad continuas, siendo la más común e importante la normal. En este capitulo se considera esta distribución y otras que se derivan de la misma.

6.2. Distribución Normal

Denominada también distribución de Gauss o distribución gaussiana, es un modelo de probabilidad de gran importancia pues permite modelar numerosos fenómenos naturales, sociales, psicológicos, entre otros. Algunos ejemplos son:

1. Caracteres morfológicos de individuos (personas, animales, plantas,etc) entre ellos: tallas, pesos, diámetros, perímetros.
2. Caracteres fisiológicos, por ejemplo: efecto de una misma dosis de un fármaco, o de una misma cantidad de abono.
3. Caracteres sociológicos, por ejemplo: consumo de cierto producto por un mismo grupo de individuos, puntuaciones de examen.
4. Caracteres psicológicos, por ejemplo: cociente intelectual, grado de adaptación a un medio,...
5. Errores cometidos al medir ciertas magnitudes.
6. Valores estadísticos muestrales, por ejemplo : la media.
7. Otras distribuciones como la binomial o la de Poisson son aproximaciones normales, ...

La distribución normal fue reconocida por primera vez por el francés Abraham de Moivre (1667-1754). Posteriormente, Carl Friedrich Gauss (1777-1855) elaboró desarrollos más profundos y formuló la ecuación de la curva, de ahí que también se le conozca, más comúnmente, como la çampana de Gauss". El termino "campana" proviene del hecho que la gráfica de su función de densidad tiene una forma acampanada y es simétrica respecto de su media.
La distribución normal aparece como el límite de varias distribuciones de probabilidad continuas y discretas, es decir otras distribuciones bajo ciertas condiciones se pueden aproximar a este modelo. Representa la distribución de mayor aplicación en estadística y muchos tests estadísticos están basados en la normalidad o, en una supuesta "normalidad".

Definición 6.1 (Distribución Normal) *Sea X una variable aleatoria de tipo continuo. Se dice que X sigue una distribución normal con media $\mu(-\infty < \mu < +\infty)$ y varianza $\sigma^2(\sigma > 0)$, lo que se denota $X \sim N(\mu, \sigma^2)$, si su función de densidad de probabilidad está dada por*

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma^2}\right)^2} ; x \in \mathbb{R} \tag{6.1}$$

donde

$\mathrm{e} = 2{,}718281828,\ \pi = 3{,}141592654$

Propiedades de la Distribución Normal

La distribución Normal tiene las siguientes propiedades:

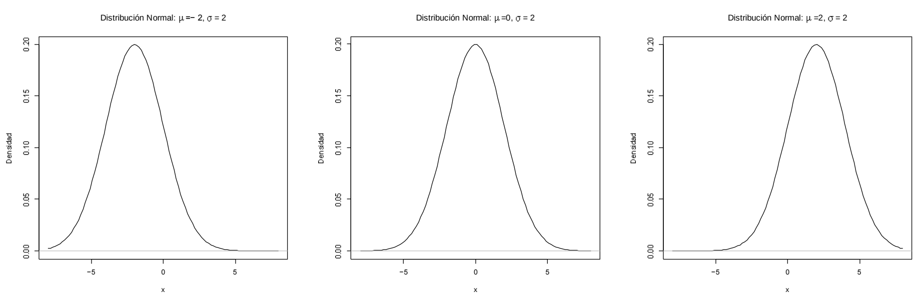
1. Es unimodal, es decir, tiene una única moda. Su valor coincide con la media y la mediana

2. Su curva es asintótica al eje de abscisas. De esta forma, puede tomar cualquier valor en el intervalo $(-\infty, +\infty)$.

3. Es simétrica respecto a su media.

4. Existe una familia de distribuciones normal, con una forma común. Cada miembro de esta familia está definido por los valores de su media y su varianza. La media indica la posición de la campana, de modo que para diferentes valores de la misma la gráfica es desplazada a lo largo del eje horizontal. La desviación estándar determina el grado de apuntamiento de la curva.

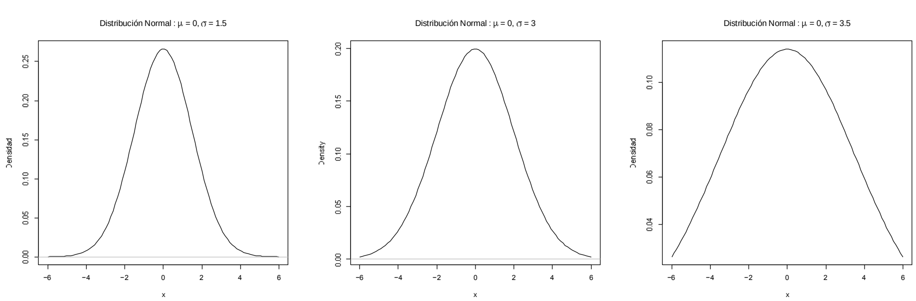
   En la figura 6.1, se muestran los gráficos de la distribución normal para varios valores de μ manteniendo fija σ^2. Aquí se aprecia que el gráfico de desplaza a la derecha a medida de que el valor de μ aumenta.



(a) $\mu = -2$ y $\sigma = 2$ (b) $\mu = 0$ y $\sigma = 2$ (c) $\mu = 2$ y $\sigma = 2$

Figura 6.1: Distribución Normal para distintos valores de μ y σ constante

Por otro lado, en la figura 2, se varía el valor de σ^2 manteniendo fijo el valor de μ, provocando esta situación que a medida que aumenta el valor de σ^2 la gráfica se va extendiendo sobre el eje X, es decir las colas se hacen mas grandes.



(a) $\mu = 0$ y $\sigma = 1,5$ (b) $\mu = 0$ y $\sigma = 2$ (c) $\mu = 2$ y $\sigma = 3,5$

Figura 6.2: Distribución Normal para distintos valores de σ y μ constante

5. El área bajo la curva comprendido entre:
   - más o menos una desviaciones estándar de la media es igual a 68,3 %.
   - más o menos dos desviaciones estándar de la media es igual a 95,4 %.
   - más o menos tres desviaciones estándar de la media es igual a 99,7 %.

Teorema 6.1 *Sea $X \sim N(\mu, \sigma^2)$. Para a y b constantes, la variable aleatoria $Y = aX + b$ sigue una distribución normal con parámetros $\mu_Y = a\mu + b$ y $\sigma_Y^2 = a^2\sigma^2$.*

Teorema 6.2 *Sean $X_1, \ldots, X_n$, n variables aleatorias independientes tales que $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, \ldots, n$. Entonces, la variable aleatoria $W = \sum_{i=1}^{n} X_i$ sigue una distribución normal con media $\sum_{i=1}^{n} \mu_i$ y varianza $\sum_{i=1}^{n} \sigma_i^2$*

Corolario 6.3 *Sean $X_1, \ldots, X_n$, n variables aleatorias independientes tales que $X_i \sim N(\mu, \sigma^2)$, $i = 1, \ldots, n$. Entonces, la variable aleatoria $W = \sum_{i=1}^{n} X_i$ sigue una distribución normal con media μ y varianza $n\sigma^2$*

De la familia de distribuciones normal, la más utilizada es la **distribución normal estándar**, la cual es una distribución normal con media 0 y varianza 1.

Definición 6.2 (Distribución Normal Estándar) *Sea Z una variable aleatoria de tipo continuo. Se dice que Z sigue una distribución normal con parámetros $\mu = 0$ y $\sigma = 1$, lo que se denota $Z \sim N(0, 1)$, si su función de densidad de probabilidad está dada por*

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}z^2}; z \in \mathbb{R} \tag{6.2}$$

Su importancia radica en que existen tablas publicadas que permiten calcular en forma sencilla la probabilidad de observar un dato menor o igual a un cierto valor z. Para determinar las áreas bajo la curva de función de densidad normal se requiere integrar la ecuación 6.1, procedimiento para el cual no existe una solución exacta.
Esta tabla, es una tabla de doble entrada y presenta la probabilidad para $Z < z$, para valores de z en el intervalo $(-3{,}9; 3{,}9)$. En la columna de la izquierda se encuentra la parte entera y el primer decimal de z, y en la fila superior las centésimas de z. En la casilla donde se interceptan la fila y la columna correspondientes, se ubica el valor de la probabilidad de que $Z < z$. Mediante el proceso inverso podemos obtener el valor z que tiene a su izquierda un área o probabilidad igual α.

Definición 6.3 (Percentil de la distribución normal estándar) *Considérese* $Z \sim N(0,1)$ *y* α *un valor en el intervalo* $(0,1)$. *Sea* z_α *el número real que satisface*

$$P(Z > z_\alpha) = \alpha \tag{6.3}$$

Entonces a z_α *se le denomina el percentil* $(1-\alpha)100$ *de la distribución normal estándar.*

De esta forma, z_α es aquel valor de la distribución normal estándar que tiene a su izquierda (derecha) un área igual a $1-\alpha$ (α). Para encontrar su valor, obsérvese que $P(Z \leq z_\alpha) = 1-\alpha$. Por tanto, usando las tablas de la distribución normal estándar se debe ubicar el valor z que acumule una probabilidad de $1-\alpha$, este valor será z_α.

Ejemplo 6.1 *Obtener el valor* Z *que tiene a su derecha un área igual a* $0{,}05$.
Este es el percentil 95 *y se denota por* $Z_{0{,}05}$. *Luego, usando la tabla de la normal estándar se tiene que* $Z_{0{,}05} = 1{,}645$.

Ejemplo 6.2 *Obtener el valor* Z *que tiene a su izquierda un área igual a* $0{,}025$.
El valor solicitado es el percentil $2{,}5$, *denotado por* $Z_{0{,}975}$. *Usando la tabla de la normal estándar se tiene que* $Z_{0{,}975} = -1{,}96$. *Obsérvese que este valor es el mismo valor que tiene a su derecha una área igual a* $0{,}025$, *pero con signo negativo, es decir,* $Z_{1-\alpha} = -Z_\alpha$. *Esto se debe a la simetría de la distribución normal alrededor de su media.*

Ejemplo 6.3 *Si* $Z \sim N(0,1)$, *obtener:*

$P(Z < -0{,}1)$, $\quad P(Z < 0{,}1)$, $\quad P(Z > 1{,}5)$

Usando la tabla de la distribución normal estándar se tiene que:

$P(Z < -0{,}1) = 0{,}4602$, $\quad P(Z < 0{,}1) = 0{,}5398$, $\quad P(Z > 1{,}5) = 0{,}0668$

Cualquier variable aleatoria X que siga una distribución $N(\mu, \sigma)$, se puede transformar en una variable Z con una distribución normal estándar, simplemente aplicando la conversión

$$Z = \frac{X-\mu}{\sigma} \tag{6.4}$$

Esta transformación recibe el nombre de **estandarización** ó **tipificación**. De acuerdo al teorema 6.1, la variable aleatoria Z sigue una distribución normal. Una de las ventajas de la estandarización es que la distribución no depende de los parámetros, pues en este caso la media siempre será cero y la varianza uno. Por tanto, la distribución es única y el gráfico de la función de densidad también.De esta forma, la estandarización resulta de especial interés en la práctica, pues como se dijo antes, existen tablas publicadas asociadas con la distribución normal estándar, lo que simplifica el calculo de probabilidades asociadas con cualquier distribución normal.

Ejemplo 6.4 *Sea $X \sim N(50, 5)$. ¿Cuál es la probabilidad de que X tome un valor entre 45 y 55?. La probabilidad que se desea calcular es $P(45 \leq X \leq 55)$. Estandarizando se tiene que*

$$\begin{aligned} P(45 \leq X \leq 55) &= P(X \leq 55) - P(X \leq 45) \\ &= P\left(Z \leq \frac{55-50}{5}\right) - P\left(Z \leq \frac{45-50}{5}\right) \\ &= P(Z \leq 1) - P(Z \leq -1) = 0{,}8413 - 0{,}1587 = 0{,}6826 \end{aligned}$$

Ejemplo 6.5 *Supóngase que el peso de los miembros de un equipo de fútbol sigue una distribución aproximadamente normal, con una media de 65 Kg y una desviación estándar de 2 Kg. ¿Cuál es la probabilidad de que jugador elegido al azar tenga un peso superior a 70 Kg?.*
Procediendo de igual manera que en el ejemplo anterior, se calcula la probabilidad solicitada.

$$\begin{aligned} P(X > 70) &= 1 - P(X \leq 70) = 1 - P\left(Z \leq \frac{70-65}{2}\right) \\ &= 1 - P(Z \leq 2{,}5)1 - 0{,}9938 = 0{,}0062 \end{aligned}$$

Ejemplo 6.6 *En el ejemplo anterior, determinar entre que valores está aproximadamente el 95 % de los pesos de los jugadores.*
Se sabe que el 95,4 % de los datos se encuentra en el intervalo $(\mu \mp 2\sigma)$. Por tanto, en el intervalo $(61, 69)$ se encuentra aproximadamente el 95 % de los datos.

6.3. Aproximación de la distribución binomial a la normal

Si n es grande la evaluación de la función de probabilidad binomial resulta difícil y poco práctico. Bajo ciertas condiciones, una distribución binomial se puede aproximar a una distribución normal con media $\mu = np$ y varianza $\sigma^2 = np(1-p)$. Mientras más grande sea n y más cercano a 0,5 esté p, mejor es la aproximación.
Si $np \geq 5$ y $n(1-p) \geq 5$, la aproximación realizada se considera buena. En caso de que se pueda aproximar, se debe tener en cuenta que estamos pasando de una variable discreta (binomial) a una continua (normal), y por tanto, son distribuciones diferentes. Lo que significa que se tienen que hacer algunos ajustes que permitan garantizar que la aproximación realizada sea lo más precisa posible. Este ajuste se denomina corrección por continuidad
Si la probabilidad solicitada es $P(X = k)$, haciendo uso de la distribución normal no se puede calcular directamente, pues como se dijo anteriormente, en una distribución continua todas estas probabilidades valen 0. En este caso la corrección por continuidad consiste en tomar un intervalo de longitud 1 alrededor de k, es decir $(k - 0{,}5, k + 0{,}5)$. Por tanto, se debe calcular $P(k - 0{,}5 \leq Y \leq k + 0{,}5)$ en lugar de $P(X = k)$. Si se pide $P(X < k)$, entonces hay que calcular $P(Y \leq k - 0{,}5)$. De esta manera se garantiza que efectivamente X sea menor estrictamente que k. Por otro lado, si se

quiere calcular $P(X \leq k)$, se debe calcular $P(Y \leq k + 0{,}5)$, ya que X debe estar incluido.

Ejemplo 6.7 *Se lanza un dado* 100 *veces. ¿ Cuál es la probabilidad de obtener un número menor o igual a* 2*: (a)* 30 *veces, (b) menos de* 40 *veces y (c) por lo menos 35 veces.*
La variable aleatoria $X = \{$ *número de veces que se obtiene un número menor o igual a* $2\}$*, sigue una distribución binomial de parámetros* $n = 100$ *y* $p = \frac{1}{3}$*. De esta forma, las probabilidades solicitadas están dadas por:*

- $P(X = 30) = \binom{100}{30}\left(\frac{1}{3}\right)^{30}\left(\frac{2}{3}\right)^{70}$.
- $P(X < 40) = P(X \leq 39) = \sum_{k \leq 39} \binom{100}{k}\left(\frac{1}{3}\right)^{k}\left(\frac{2}{3}\right)^{100-k}$.
- $P(X \geq 35) = \sum_{k \geq 35} \binom{100}{k}\left(\frac{1}{3}\right)^{k}\left(\frac{2}{3}\right)^{100-k}$.

Obsérvese que en todos los casos, los cálculos no son fáciles de realizar. Dado que $np = 33{,}33 > 5$ *y* $n(1-p) = 66{,}67 > 5$*, se puede hacer uso de la aproximación normal. Esto es, usando una variable* Y *tal que* $Y \sim N(33{,}33, 4{,}71)$*, se hace el calculo de las probabilidades solicitadas.*

$$\begin{aligned} P(X = 30) &= P(29{,}5 \leq Y \leq 30{,}5) = P(Y \leq 30{,}5) - P(Y \leq 29{,}5) \\ &= P(Z \leq -0{,}6) - P(Z \leq -0{,}81) = 0{,}0653 \end{aligned}$$

$$P(X < 40) = P(Y \leq 39{,}5) = P(Z \leq 1{,}31) = 0{,}9049$$

$$\begin{aligned} P(X \geq 35) &= 1 - P(X < 35) = 1 - P(Y \leq 34{,}5) = 1 - P(Z \leq 0{,}248) \\ &= 0{,}598 \end{aligned}$$

Ejemplo 6.8 *El* 25 % *de los habitantes de la ciudad votan por cierto partido político. Se encuesta a* 100 *personas. Sea* X *la variable aleatoria que representa el número de personas que votan por dicho partido. ¿Cuál es la distribución de* X*?. Calcular la probabilidad de que entre las personas encuestadas haya entre* 35 *y* 40 *simpatizantes de dicha organización política.*
Las personas votan o no por el partido. Por tanto, $X \sim B(100, 0{,}25)$*. Dado que* $np = 25 > 5$ *y* $n(1-p) = 75 > 5$*, la probabilidad solicitada se puede calcular usando la aproximación normal. Sea*

$Y \sim N(25, 8{,}66)$. *Luego*

$$
\begin{aligned}
P(35 \leq X \leq 40) &= P(34{,}5 \leq Y \leq 39{,}5) \\
&= P(Y \leq 39{,}5) - P(Y \leq 34{,}5) \\
&= P(Z \leq 1{,}67) - P(Z \leq 1{,}09) = 0{,}0904
\end{aligned}
$$

6.4. Aproximación de la distribución poisson a la normal

Cuando la media de una distribución poisson es relativamente grande, puede aproximarse a una distribución normal con parámetros $\mu = \lambda$ y $\sigma = \sqrt{\lambda}$. Se considera que esta aproximación es buena cuando $\lambda > 15$ y se considera aceptable para $\lambda \geq 10$.
Igual que en el caso de la aproximación de la distribución binomial a la normal, aquí se debe hacer la corrección por continuidad.

Ejemplo 6.9 *Sea* $X \sim P(30)$. *Obtener* $P(25 \leq X \leq 32)$.
Para calcular $P(25 \leq X \leq 32)$ *usamos la aproximación normal*

$$
\begin{aligned}
P(25 \leq X \leq 32) &= P(24{,}5 \leq Y \leq 32{,}5) \\
&= P(Y \leq 32{,}5) - P(Y \leq 24{,}5) \\
&= P(Z \leq 0{,}46) - P(Z \leq -1) = 0{,}5185
\end{aligned}
$$

Ejemplo 6.10 *Si un banco procesa en promedio* 20 *cheques al día, cuál es la probabilidad de que reciba:*

1. *15 cheques en un día dado.*

2. *entre 18 y 24 cheques en un día dado.*

La variable $X = \{$*número de cheques procesados por día*$\}$, *es una variable aleatoria poisson con parámetro* $\lambda = 20$. *Obsérvese que* $\lambda > 15$, *es decir, usar la aproximación normal es una buena alternativa. Sea* $Y \sim N(20, 4{,}47)$, *entonces*

$$
\begin{aligned}
P(X = 15) &= P(14{,}5 \leq Y \leq 15{,}5) \\
&= P(Z \leq -1{,}01) - P(Z \leq -1{,}23) \\
&= 0{,}1562 - 0{,}1093 = 0{,}0469
\end{aligned}
$$

$$\begin{aligned} P(18 \leq X \leq 24) &= P(17{,}5 \leq Y \leq 24{,}5) \\ &= P(Z \leq 1{,}01) - P(Y \leq -0{,}56) \\ &= 0{,}8438 - 0{,}2877 = 0{,}5561 \end{aligned}$$

6.5. Distribuciones derivadas de la normal

En esta sección se muestran varios resultados importantes que indican como realizando diferentes transformaciones de la v.a. normal se pueden obtener otras variables aleatorias que resultan esenciales en la estimación por intervalos y las pruebas de hipótesis. Aquí se definirán las distribuciones de probabilidad Chi-cuadrado, t de Student y F de Fisher.

6.5.1. Distribución Chi-Cuadrado

Teorema 6.4 (Distribución Chi-Cuadrado (χ^2)) *Si $Z_1, \ldots, Z_n$ son n variables aleatorias independientes e idénticamente distribuidas según la normal estándar, entonces la variable aleatoria* $Y = \sum_{i=1}^{n} Z_i^2$ *, sigue una distribución Chi-cuadrado con n grados de libertad, lo que se denota* $Y \sim \chi_n^2$*.*

El número de **grados de libertad** para un conjunto de datos simple es el número de valores muestrales que pueden variar después de que ciertas restricciones hayan sido impuestas a todos los valores de la data.
La distribución χ^2 es la distribución a la que se ajusta la distribución muestral de la varianza de una población normal. Entre sus **aplicaciones** se encuentran:

- Inferencia sobre la varianza
- Test de bondad de ajuste
- Test de independencia
- Test de homogeneidad

Las **características** de la Distribución Chi Cuadrado son:

1. Es de tipo continuo.
2. Es asimétrica positiva.
3. Su rango es el intervalo $(0, \infty)$.
4. Tiene un solo parámetro, sus grados de libertad.
5. Su media es n y su varianza $2n$.

6. Para cada valor de n hay una distribución χ^2 diferente. Esto es, existe una familia de distribuciones χ^2.
7. Cuando n es suficientemente grande, se puede aproximar por una distribución normal.

De acuerdo con las características antes señaladas la gráfica de la distribución χ^2 se muestra en la figura 6.3

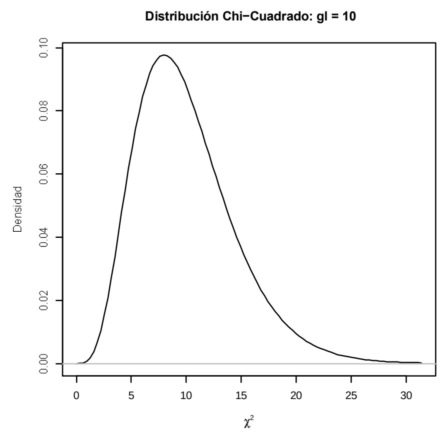


Figura 6.3: Distribución Chi-Cuadrado

Corolario 6.5 *Si $X_1, \ldots, X_n$ son n variables aleatorias independientes tales que $X_i \sim \chi^2_{\nu_i}$, $i = 1, \ldots, n$, entonces*

$$S = \sum_{i=1}^{n} X_i \sim \chi^2_{\nu} \tag{6.5}$$

donde $\nu = \sum_{i=1}^{n} \nu_i$

Definición 6.4 (Percentil de la distribución χ^2) *Si $X \sim \chi^2_n$ y α un valor en el intervalo $(0, 1)$. Entonces $\chi^2_{n;\alpha}$ es el número real que satisface*

$$P(X > \chi^2_{n;\alpha}) = \alpha \tag{6.6}$$

Entonces a $\chi^2_{n;\alpha}$ se le denomina el percentil $(1-\alpha)100$ de la distribución Chi-cuadrado con n grados de libertad. Su significado es el mismo que en el caso de la distribución normal.

Igual que en el caso de la distribución normal, existen tablas publicadas asociadas con la distribución χ^2, para diferentes valores de n. Esto simplifica el calculo de probabilidades asociadas con la

distribución χ^2.
Como ocurre con la distribución normal, la función de densidad chi-cuadrado no es nada manejable. Es por ello que en vez de trabajar con su expresión, es preferible trabajar con sus tablas. Estas tablas pueden informarnos bien del propio valor de la función de distribución, o su complemento. La primera columna no da los grados de libertad y la primera fila la probabilidad que deja a su derecha el punto que nos indica la parte central de la tabla. Esto es, por ejemplo, el valor 5,22935 correspondiente a una chi-cuadrado de 15 grados de libertad deja a su derecha una probabilidad o área igual a 0,99. igualmente, esta tabla permite ubicar aquel valor que tiene a su derecha un área o probabilidad determinada. Por ejemplo, el valor de una chi-cuadrado de 15 grados que tiene a su derecha un área igual 0,05 es 24,9958. Si se quiere el valor que tiene a su izquierda una probabilidad dada α, simplemente se ubica en la tabla aquel valor que tiene a su derecha un área igual a $1-\alpha$.
Por otro lado, la probabilidad de que X sea menor a 5,22935 es 0,01

Ejemplo 6.11 *Obtener el valor de χ^2_{15} que tiene a su derecha un área igual a 0,01.*
Este es el percentil 99 y se denota por $\chi^2_{15;0,01}$. Usando la tabla de la distribución χ^2 se tiene que $\chi^2_{15;0,01} = 30{,}5779$

Ejemplo 6.12 *Obtener el valor de χ^2_{15} que tiene a su izquierda un área igual a 0,05.*
Este es el percentil 5 y se denota por $\chi^2_{15;0,95}$. Usando la tabla de la distribución χ^2 se tiene que $\chi^2_{15;0,95} = 7{,}26094$

Ejemplo 6.13 *Si $X \sim \chi^2_{10}$, obtener:*

$P(X < 2{,}558)$, $\quad P(X > 19{,}6751)$

Usando la tabla de la distribución Chi-cuadrado se tiene que:

$P(X < 2{,}558) = 0{,}01$, $\quad P(X > 18{,}3070) = 0{,}05$

Teorema 6.6 (Distribución t-student) *Sean $X \sim N(0,1)$ y $Y \sim \chi^2_n$, variables aleatorias independientes. Entonces, la variable aleatoria*

$$T = \frac{X}{\sqrt{\frac{Y}{n}}} \tag{6.7}$$

sigue una distribución t de student con n grados de libertad, lo que se denota $T \sim t_n$.

En general, no se dispone de la desviación standard de la población, sino de una estimación calculada a partir de una muestra extraída de la misma. Si la población sigue una distribución normal y el tamaño de muestra es pequeño, una alternativa a la distribución normal estándar es la distribución t de student.
Entre sus **aplicaciones** se encuentra la inferencia estadística acerca de una y dos medias poblacionales. Sus características son:

1. Es de tipo continuo.
2. Es simétrica.
3. Su rango es el intervalo $(-\infty, +\infty)$.
4. Tiene un solo parámetro, sus grados de libertad.
5. Su media es cero y su varianza $\frac{n}{(n-2)}$.
6. Para cada valor de n hay una distribución t diferente.
7. A medida que n crece, la distribución t se aproxima a la normal. Esto es, cuando n es lo suficientemente grande, se puede aproximar por una distribución normal.

De acuerdo con las características antes señaladas la gráfica de la distribución t se muestra en la figura 6.4

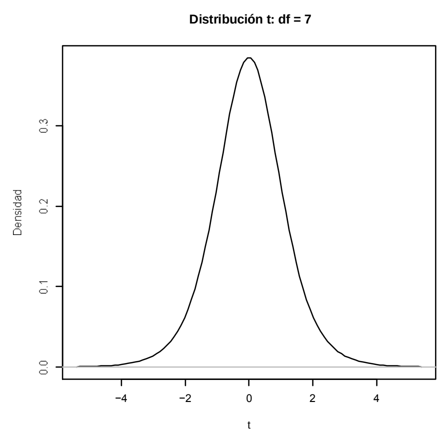


Figura 6.4: Distribución t-student

Definición 6.5 (Percentil de la distribución t) *Si $X \sim t_n$ y α un valor en el intervalo $(0, 1)$. Entonces $t_{n;\alpha}$ es el número real que satisface*

$$P(X > t_{n;\alpha}) = \alpha \tag{6.8}$$

Entonces a $t_{n;\alpha}$ se le denomina el percentil $(1-\alpha)100$ de la distribución t con n grados de libertad. Su significado es el mismo que en el caso de la distribuciones normal y χ^2.

Tal y como se hizo con las distribuciones chi-cuadrado y normal, tanto para el cálculo de probabilidades como para el calculo de los percentiles, en la distribución t-Student se utilizarán tablas construidas para facilitar la tarea.
En la primera fila de estas tablas se muestran los valores de α, es decir, la probabilidad de que la variable tome un valor mayor que el considerado. En la primera columna se muestran los grados de libertad y en el centro de la tabla nos da los valores de la variable. De esta forma, por ejemplo, una t-Student de 15 grados de libertad deja a la derecha del punto 2,602 una probabilidad igual a 0,01. Es importante hacer notar que dada la simetría de esta distribución, una t-Student de 15 grados de libertad deja a la izquierda del punto $-2{,}602$ una probabilidad igual a 0,01. Igual que en los casos anteriores, esta tabla permite ubicar el valor que tiene a su derecha un área α determinada. El valor una t-Student de 20 grados de libertad que tiene a su derecha una probabilidad igual a 0,025 es 2,086. La probabilidad de que X sea menor a 2,086 es 0,975.

Ejemplo 6.14 *Obtener el valor de t_5 que tiene a su derecha un área igual a* 0,05.
El valor solicitado es el percentil 95*,* $t_{5;0,05}$*. Usando la tabla de la distribución t, se tiene que* $t_{5;0,05} = 2{,}015$.

Ejemplo 6.15 *Obtener el valor de t_5 que tiene a su izquierda un área igual a* 0,01.
Este valor es el percentil 1*, es decir,* $t_{5;0,99}$*. Dada la simetría de la distribución t, este valor es equivalente a* $-t_{5;0,01} = -3{,}365$

Ejemplo 6.16 *Si $X \sim t_{10}$, obtener*

$P(X < 1{,}812)$, $\quad P(X < -3{,}169)$, $\quad P(X > -0{,}25)$

Usando la tabla de la distribución t se tiene que:

$P(X < 1{,}812) = 0{,}05$, $\quad P(X < -3{,}169) = 0{,}005$,$P(X > -0{,}25) = 0{,}6$

Teorema 6.7 (Distribución F-Snedecor) *Sean $V \sim \chi^2_n$ y $W \sim \chi^2_m$, variables aleatorias independientes. Entonces, la variable aleatoria*

$$X = \frac{\frac{V}{n}}{\frac{W}{m}} \tag{6.9}$$

sigue una distribución F con n y m grados de libertad, lo que se denota $X \sim F_{n,m}$.

Sus principales usos son los de la contrastación de la igualdad de varianzas de dos poblaciones normales y, fundamentalmente, el análisis de la varianza y el diseño de experimentos, técnicas que permiten detectar la existencia de diferencias significativas entre muestras diferentes.

Las características de la distribución F son:

1. Es de tipo continuo.
2. Es asimétrica positiva.
3. Su rango es el intervalo .
4. Tiene dos parámetros, n y m.
5. Su media es $\frac{m}{m-2}$ y su varianza $\frac{2m^2(n+m-2)}{n(m-2)(m-4)}$.
6. Para cada par de valores de n y m, hay una distribución F diferente.
7. A medida que n y m aumentan, la distribución F se aproxima a la normal.

De acuerdo con las características antes señaladas la gráfica de la distribución F se muestra en la figura 6.5

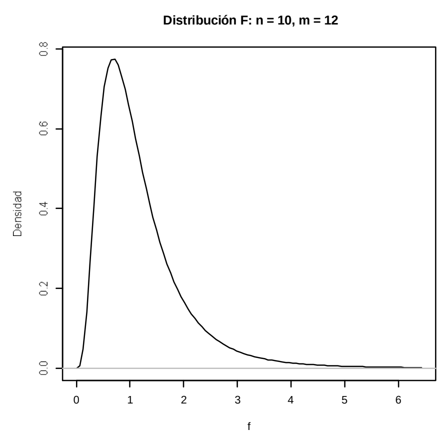


Figura 6.5: Distribución F

Definición 6.6 (Percentil de la distribución F) *Si $X \sim F_{n,m}$ y α un valor en el intervalo $(0,1)$. Entonces $F_{n,m;\alpha}$ es el número real que satisface*

$$P(X > F_{n,m;\alpha}) = \alpha \tag{6.10}$$

A $F_{n,m;\alpha}$ se le denomina el percentil $(1-\alpha)100$ de la distribución F con n y m grados de libertad.

Otra propiedad de la distribución F es que la misma satisface la igualdad

$$F_{n,m;\alpha} = \frac{1}{F_{m,n;1-\alpha}} \tag{6.11}$$

Para esta distribución también se han elaborado tablas que permiten fácilmente calcular probabilidades y percentiles. En la tabla de la F de Fisher-Snedecor se presentan en la primera fila α y en la segunda los grados de libertad del numerador. En la primera columna, se presentan los grados de libertad del denominador. La tabla ofrece la probabilidad o área que queda a la derecha del punto seleccionado. De esta forma, en el interior de la tabla se muestran los valores que dejan a su derecha una probabilidad α para los grados de libertad n y m seleccionados. Por ejemplo, el valor 5,61 de una distribución F de 7 y 9 grados de libertad deja a su derecha una probabilidad igual a 0,01. La probabilidad de que X sea menor a 5,61 es, por tanto, 0,99.

Ejemplo 6.17 *Obtener el valor de $F_{5,10}$ que tiene a su derecha un área igual a 0,05.*
El valor solicitado es el percentil 95, $F_{5,10;0,05}$. Usando la tabla de la distribución F, se tiene que $F_{5,10;0,05} = 3{,}33$.

Ejemplo 6.18 *Obtener el valor de $F_{5,10}$ que tiene a su izquierda un área igual a 0,025.*
Este valor es el percentil 2,5, es decir, $F_{5,10;0,975}$. Usando la propiedad de reciprocidad de la distribución F, $F_{5,10;0,975} = \frac{1}{F_{10,5;0,025}}$. Luego, $F_{5,10;0,975} = \frac{1}{6{,}62} = 0{,}1511$

Ejemplo 6.19 *Si $X \sim F_{10,10}$, calcular*

$P(X < 4{,}85), \qquad P(X > 3{,}98)$

Usando la tabla de la distribución F se tiene que:

$P(X < 4{,}85) = 0{,}99, \qquad P(X > 3{,}98)) = 0{,}05$

6.6. Ejercicios

1. Si $Z \sim N(0,1)$, obtener
   a) $P(Z < -1{,}96)$
   b) $P(Z < 1{,}65)$
   c) $P(Z < 2{,}33)$
   d) $P(Z < -2{,}33)$
   e) $P(Z > 1{,}96)$
   f) $P(-1{,}96 < Z < 1{,}96)$
   g) $P(-4 < Z < 4)$
2. Si $Z \sim N(0,1)$, obtener:
   a) el valor z que tiene a su izquierda una probabilidad o área igual a 0,017.

   b) el valor z que tiene a su derecha una probabilidad o área igual a 0,025.
   c) los valores z_1 y z_2 que tienen a su extremos una probabilidad o área total igual a 0,01.
   d) los valores z_1 y z_2 que tienen a su extremos una probabilidad o área total igual a 0,05.

3. Sea $X \sim N(20, 5)$. Obtener a probabilidad de que X:

   a) sea mayor a 22.
   b) sea menor a 17.
   c) este entre 16 y 24.

4. En el ejercicio 3, obtener la distribución de probabilidad de $Y = 2X + 3$ y calcular la probabilidad de que Y:

   a) sea mayor a 25.
   b) sea menor a 30.
   c) entre 22 y 34.

5. Las calificaciones de los 500 aspirantes que presentaron la última prueba de actitud académica en FACES, se distribuye normalmente con media 58 y varianza 4.

   a) Calcule la probabilidad de que un aspirante obtenga más de 50 puntos.
   b) Determine la proporción de aspirantes con calificaciones inferiores a 5 puntos.
   c) Calcule la probabilidad de que un aspirante obtenga una calificación entre 50 y 60 puntos.
   d) ¿Cuántos aspirantes obtuvieron calificaciones comprendidas entre 50 y 60 puntos?.

6. El 5 % de los componentes elaborados por una máquina resultan defectuosos. Si se selecciona un lote de 500 componentes, ¿cual es la probabilidad de que hayan menos de 30 defectuosos?.

7. Una moneda balanceada se lanza 400 veces. Determinar la probabilidad de que el número de caras:

   a) sea superior a 150.
   b) sea inferior a 250.
   c) este entre 175 y 300.
   d) este entre 150 y 250.

8. Un futbolista acierta en el arco contrario (hace gol) en el 70 % de los juegos en los que participa. Si en un campeonato en particular de 38 juegos, cual es la probabilidad de que haga gol en más de 15 juegos.

9. Un vendedor de pólizas de seguro atiende a 100 clientes de los cuales sólo el 15 % adquiere dicha póliza. ¿Cuál es la probabilidad de que un día cualquiera, venda por lo menos 20 pólizas?.

10. A un consultorio médico acuden en promedio 25 pacientes al día. Calcular la probabilidad de que un día cualquiera, acudan más de 15 pacientes.

11. El número de bombillas que fallan antes de cumplir 1000 horas de funcionamiento, es una variable aleatoria poisson con $\lambda = 20$. ¿Cuál es la probabilidad de que en 1000 horas fallen:

    a) exactamente 15 bombillas.
    b) a lo sumo 15 bombillas.
    c) por lo menos 15 bombillas.
    d) entre 15 y 25 bombillas.

12. Sea $X \sim \chi^2_{10}$. Obtener la probabilidad de que X:

    a) sea mayor a 23,2093.
    b) sea menor a 2,55821.
    c) este entre 3,24697 y 20,4832.

13. Sea $X \sim \chi^2_{28}$. Obtener:

    a) el valor x que tiene a su derecha una probabilidad o área igual a 0,05.
    b) el valor x que tiene a su izquierda una probabilidad o área igual a 0,05.
    c) los valores x_1 y x_2 que tienen a sus extremos una probabilidad o área igual a 0,05.
    d) los valores x_1 y x_2 que tienen a sus extremos una probabilidad o área igual a 0,10.
    e) los valores x_1 y x_2 que tienen a sus extremos una probabilidad o área igual a 0,01.

14. Sea $X \sim \chi^2_{12}$. Obtener el valor x tal que:

    a) $P(X > x) = 0{,}1$.
    b) $P(X < x) = 0{,}975$.

15. En el ejercicio 14, obtener:

    a) $P(X > 26{,}2179)$.
    b) $P(X < 5{,}222603)$.
    c) $P(4{,}40379 < X < 23{,}3367)$.

16. Sea $X \sim t_{15}$. Obtener la probabilidad de que X:

    a) sea mayor a 1,341.
    b) sea menor a 2,602.
    c) este entre $-2{,}131$ y 1,753.

17. Sea $X \sim t_{20}$. Obtener:

    a) el valor x que tiene a su derecha una probabilidad o área igual a 0,05.

    b) el valor x que tiene a su izquierda una probabilidad o área igual a 0,05.

    c) los valores x_1 y x_2 que tienen a sus extremos una probabilidad o área igual a 0,05.

    d) los valores x_1 y x_2 que tienen a sus extremos una probabilidad o área igual a 0,10.

    e) los valores x_1 y x_2 que tienen a sus extremos una probabilidad o área igual a 0,01.

18. Si $X \sim t_{12}$, obtener el valor x tal que:

    a) $P(X > x) = 0{,}05$.

    b) $P(X < x) = 0{,}025$.

19. En el ejercicio 18, obtener:

    a) $P(X > 2{,}179)$.

    b) $P(X < 3{,}055)$.

    c) $P(-1{,}356 < X < 2{,}179)$.

    d) $P(-2{,}179 < X < 2{,}179)$.

20. Sea $X \sim F_{15;20}$. Obtener la probabilidad de que X:

    a) sea inferior a 3,09.

    b) sea superior a 2,20.

21. En el ejercicio 20, obtener obtener el valor x tal que:

    a) $P(X > x) = 0{,}975$

    b) $P(X < x) = 0{,}01$

    c) $P(X < x) = 0{,}90$

    d) $P(X > x) = 0{,}05$