Cálculo de valores y probabilidades de Distribuciones Muestrales

Giovanni Sanabria Brenes

Resumen

Se definen las distribuciones muestrales necesarias para el curso. Se aborda el cálculo de valores y probabilidades para cada distribución, utilizando el app Probability Distribucions.

1 App Probability Distribucions

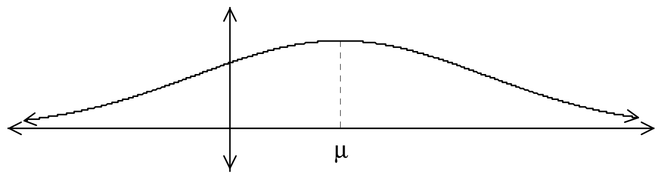
El app Probability Distribucions fue desarrollado por el Departamento de Estadística y Ciencias Actuariales de la Universidad de Iowa:

Este app permite: el cálculo de probabilidades de una determina distribución, calculo de valores que acotan una probabilidad acumulada a la derecha o a la izquierda para una determinada distribución y representar gráficamente la función de distribución para la distribución binomial, geométrica, Poisson, hipergeométrica, y las distribuciones binomial negativa. probabilidades de cómputo, determinar los percentiles, y el argumento de la función de densidad de probabilidad para la normal (Gaussiana), t, chi-cuadrado, F, exponencial, gamma, beta y distribuciones normales de registro. El app está disponible para android y ios.

2 Distribución normal

Recuerde que si $X \sim N\left(\mu, \sigma^2\right)$ entonces $Z = \dfrac{X-\mu}{\sigma} \sim N(0,1)$, es decir, Z sigue una distribución normal estándar. La función acumulada de la distribución normal estándar se denota por

$$F_Z(k) = P(Z \leq k) = \phi(k)$$



Ejemplo 1 *Sea* $Z \sim N(0,1)$ *Para obtener* $P(Z < -1.52)$ *utilizando el app Probability Distribucions, se selecciona la distribución Normal (por defecto sale la normal estándar:* $\mu = 0$ *y* $\sigma = 1$*), se escribe en la casilla correspondiente:*

$$x = -1.52$$

y se selecciona la opción $P(X < x)$, *obteniendo que*

$$P(Z < -1.52) = 0.06426$$

y por simetría se tiene que

$$P(Z > 1.52) = P(Z < -1.52) = 0.06426$$

Además

$$P(Z > -1.52) = 1 - P(Z < -1.52) = 1 - 0.06426 = 0.93574$$

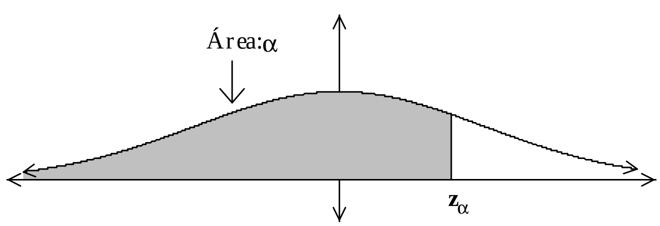
y

$$P(Z < 1.52) = 1 - P(Z > 1.52) = 1 - 0.06426 = 0.93574$$

Esta distribución y las que veremos más adelante son continuas, por lo tanto, no hay probabilidad puntual, por ejemplo: $P(Z > 1.52) = P(Z \geq 1.52)$. Por eso no se utilizó el igual al calcular probabilidades de complementos, en el ejemplo anterior.

Definición 1 *Si* $Z \sim N(0,1)$ *entonces se define* ***el valor z:***

$$z_\alpha : \textit{valor de } Z \textit{ que acumula un área a la izquierda de } \alpha\%.$$



Es decir $P(Z \leq z_\alpha) = \alpha$.

Note que si α es menor a 0.5, entonces z_α es negativo y viceversa. De igual forma, si α es mayor a 0.5, entonces z_α es positivo y viceversa.

Teorema 1 *Se tiene que* $z_\alpha = -z_{1-\alpha}$

Ejemplo 2 *Calcule* $z_{0.43}$.

Utilizando el app Probability Distribucions, se selecciona la distribución Normal (debe aparecer $\mu = 0$ *y* $\sigma = 1$*) y se escribe en la casilla correspondiente:*

$$P(X < x) = 0.43$$

y se obtiene que $z_{0.43} = -0.17637$

Ejercicio 1 *Verifique que*

$$z_{0.1} \approx -1.28, \quad z_{0.05} \approx -1.645, \quad z_{0.025} \approx -1.96$$

y por simétria se tiene que

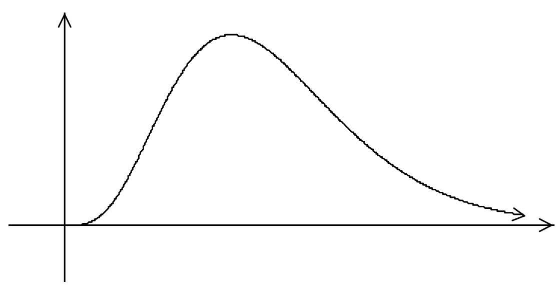
$$z_{0.9} = -z_{0.1} \approx 1.28, \quad z_{0.95} = -z_{0.05} \approx 1.645, \quad z_{0.975} = -z_{0.025} \approx 1.96$$

3 Distribución Chi Cuadrado

Definición 2 *Sea* X *una v.a.c. Se dice que* X *sigue una distribución chi cuadrado con* v *grados de libertad si y solo si*

$$X \sim Gamma(v/2, 2)$$

Se denota $X \sim \chi^2(v)$



Los grados de libertad indican las piezas de información independientes en el cálculo de X.

Teorema 2 *Sea* X *una v.a.c. tal que* $X \sim \chi^2(v)$, *entonces:*

1. $E(X) = v$

2. $Var(X) = 2v$

Definición 3 *Si* χ^2 *tiene una distribución chi cuadrado con* v *grados de libertad entonces se define el* ***valor chi*** *por*

$$\chi^2_{\alpha,v} : \text{valor de } \chi^2 \text{ que acumula un área a la izquierda de } \alpha\%.$$

Es decir $P\left(\chi^2 \le \chi^2_{\alpha,v}\right) = \alpha$.

Para determinar el valor de $\chi^2_{\alpha,v}$ se suelen utilizar tablas.

Ejemplo 3 *Calcular el valor de* $\chi^2_{0.2,14}$.

Utilizando el app Probability Distribucions, se selecciona la distribución Chi-Square y se escribe en las casillas correspondientes:

$$v = 14, \quad P(X < x) = 0.2$$

y se obtiene que $\chi^2_{0.2,14} = 9.46733$

Ejemplo 4 *Si* $\chi^2 \sim \chi^2(10)$, *determine aproximadamente la probabilidad de que* $X > 4$.

Utilizando el app Probability Distribucions, se selecciona la distribución Chi-Square, se escribe en las casillas correspondientes:

$$v = 10, \quad x = 0.2$$

y se selecciona la opción $P(X > x)$, *obteniendo que*

$$P\left(\chi^2 > 4\right) \approx 0.94735$$

Para efectos del curso, con un valor aproximado es suficiente para tomar desiciones. De hecho si se utilizan las tablas estadísticas de la distribución χ^2, *más que una aproximación se obtiene una acotación de la probabilidad:*

$$P\left(\chi^2 > 4\right) \in \left]0.9, 0.95\right[,$$

lo cual es suficiente para hacer inferencias estadísticas. Note que el valor aproximado obtenido se encuentra en el intervalo dado por las tablas.

En el libro de texto, puede ver con detalle como se utilizan las tablas para el cálculo de valores y de probabilidades de las diferentes distribuciones a utilizar. Sin embargo, para el curso se utilizará el app, en lugar de las tablas.

El siguiente teorema indica la utilidad de esta distribución Chi Cuadrado.

Teorema 3 *Considere la población dada por la variable aleatoria* X *que sigue una distribución normal con variancia poblacional* σ^2. *Dada una muestra aleatoria de esta población* $(X_1, X_2, X_3, ..., X_n)$. *Entonces la variable*

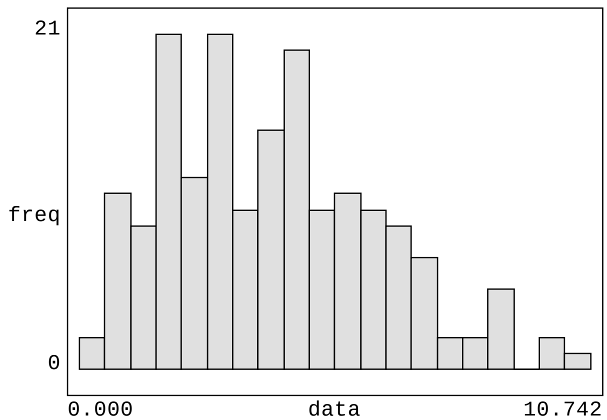
$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

sigue una distribución χ^2 *con* $v = n - 1$ *grados de libertad.*

Ejemplo 5 *Los siguientes datos son valores de una variable* Z *que sigue una distribución normal estándar:*

0.543551593	0.260990164	−0.514336612	−2.200284197
0.853214429	−0.559109129	−0.993302277	−1.477762554
−0.682542739	−0.24714637	0.693164286	0.556075608
−0.174740261	0.883016138	−0.053566134	1.07158394
0.411039337	−0.105383202	0.836494738	−1.733486719

A partir de estos valores se tomaron 180 muestras aleatorias de tamaño 6 y se registró el valor $\chi^2 = \dfrac{(n-1)S^2}{\sigma^2} = 5S^2$ *para cada muestra.*
Realizando un histograma con los valores registrados se obtiene:



Note que la distribución Chi-Cuadrado se ajusta bien a este histograma.

4 Distribución t

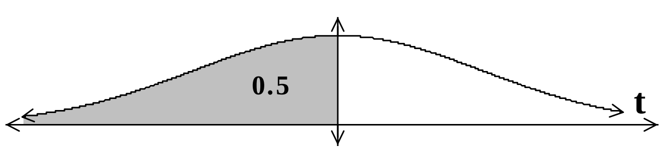
Definición 4 *Se define la función Gamma por* $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$

Definición 5 *Sea T una v.a.c. Se dice que T sigue una distribución t con v grados de libertad si y solo si su función de distribución está dada por*

$$f_T(t) = \frac{\Gamma\left(\dfrac{v+1}{2}\right)}{\Gamma\left(\dfrac{v}{2}\right)\sqrt{\pi v}}\left(1+\frac{t^2}{v}\right)^{-(v+1)/2}$$

Se denota $T \sim t(v)$.

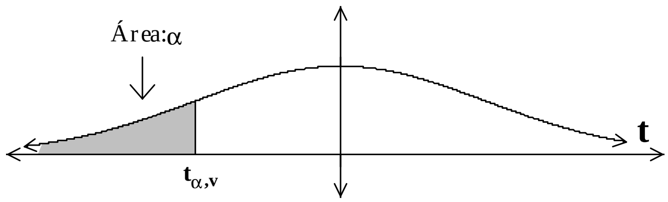
La distribución t es simétrica con respecto a su media que es 0.



Definición 6 *Si T tiene una distribución t con v grados de libertad entonces se define* ***el valor t****:*

$t_{\alpha,v}$: *valor de T que acumula un área a la izquierda de* $\alpha\%$.

Es decir $P(T \leq t_{\alpha,v}) = \alpha$



y por la simetría de t : $P(T > -t_{\alpha,v}) = \alpha$.

Teorema 4 *Dado que* $t_{\alpha,v} = -t_{1-\alpha,v}$ *entonces* $|t_{\alpha,v}| = |t_{1-\alpha,v}|$.

Ejemplo 6 *Calcular el valor de* $t_{0.975,14}$.

Utilizando el app Probability Distribucions, se selecciona la distribución t *y se escribe en las casillas correspondientes:*

$$v = 14, \quad P(X < x) = 0.975$$

y se obtiene que $t_{0.975,14} = 2.14479$.

Ejercicio 2 *Si* $t \sim t(48)$ *determine el valor de* a *que cumple que*
$0.04 = P(t > a)$ *R/* 1.78854

Ejercicio 3 *Si* $t \sim t(12)$ *determine aproximadamente*

1. $P(t > 2.2)$ *R/* 0.025
2. $P(t < -1.4)$ *R/* 0.1
3. $P(t < 1.8)$ *R/* 0.95

El siguiente relaciona las distribuciones estudiadas.

Teorema 5 *Sean* Z *y* V *variables aleatorias continuas independientes tales que*

$$Z \sim N(0,1) \quad y \quad V \sim \chi^2(v)$$

Considere la v.a.c. dada por

$$T = \frac{Z}{\sqrt{V/v}}$$

Se tiene que T *sigue una distribución* t *con* v *grados de libertad.*

El resultado siguiente presenta la utilidad de la distribución t.

Teorema 6 *Considere la población dada por la variable aleatoria* X *que sigue una distribución normal con media poblacional* μ *y varianza poblacional* σ^2. *Dada una muestra aleatoria de esta población* $(X_1, X_2, X_3, ..., X_n)$. *Entonces la variable*

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

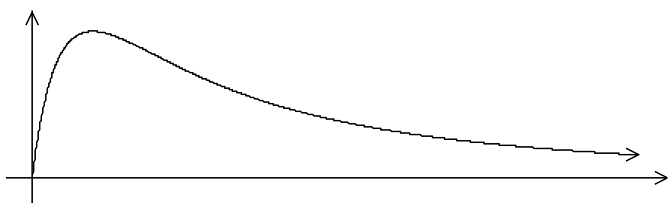
sigue una distribución t *con* $v = n - 1$ *grados de libertad.*

5 Distribución F

Definición 7 *Sea F una v.a.c. Se dice que F sigue una distribución f con v_1 y v_2 grados de libertad si y solo si su función de distribución está dada por*

$$f_F(x) = \begin{cases} \dfrac{\Gamma\left(\dfrac{v_1+v_{21}}{2}\right)\left(\dfrac{v_1}{v_2}\right)^{v_1/2} x^{v_1/2-1}}{\Gamma\left(\dfrac{v_1}{2}\right)\Gamma\left(\dfrac{v_2}{2}\right)\left(1+\dfrac{v_1 x}{v_2}\right)^{(v_1+v_2)/2}} & si \quad x > 0 \\ 0 & si \quad x \leq 0 \end{cases}$$

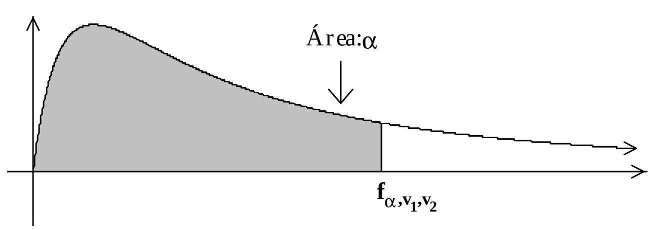
La distribución F es asimétrica:



Definición 8 *Si F tiene una distribución F con v_1 y v_2 grados de libertad entonces se define* ***el valor f****:*

$$f_{\alpha,v_1,v_2} : \text{valor de } F \text{ que acumula un área a la izquierda de } \alpha\%.$$

Es decir $P(F \leq f_{\alpha,v_1,v_2}) = \alpha$



Teorema 7 *Se tiene que*

$$f_{1-\alpha,v_1,v_2} = \frac{1}{f_{\alpha,v_2,v_1}}$$

Ejemplo 7 *Determine el valor de $f_{0.95,6,10}$ y de $f_{0.05,10,6}$.*

Utilizando el app Probability Distribucions, se selecciona la distribución F y se escribe en las casillas correspondientes:

$$d_1 = 6, \quad d_2 = 10, \quad P(X < x) = 0.95$$

y se obtiene que $f_{0.95,6,10} = 3.21717$. Por lo tanto, $f_{0.05,10,6} = \dfrac{1}{3.21717} \approx 0.310\,832$.

Ejemplo 8 *Si* $F \sim f(14, 20)$ *Acote el valor de* $P(F < 2.5)$ *y de* $P(F > 2.5)$.

Se busca el valor de α *que cumple que* $f_{\alpha,14,20} = 2.5$. *Utilizando el app Probability Distribucions, se selecciona la distribución* F, *se escribe en las casillas correspondientes:*

$$d_1 = 14, \quad d_2 = 20, \quad x = 2.5$$

y se selecciona la opción $P(X < x)$, *obteniendo que*

$$P(F < 2.5) \approx 0.96989 \in]0.95, 0.975[$$

Entonces

$$P(F > 2.5) = 1 - P(F < 2.5) \in]0.025, 0.05[$$

Ejercicio 4 *Si* $F \sim f(14, 20)$ *Acote el valor de* $P(F > 1.289)$. *R/* $]0.1, 0.9[$.

Ejercicio 5 *Si* $F \sim f(10, 16)$ *Determine aproximadamente el valor de* $P(F < 0.3)$. *R/* 0.02937

Ejercicio 6 *Suponga que* $F \sim f(8, 20)$. *Acote lo mejor posible, utilizando las tablas, el valor de* $P(F > 0.2)$. *R/* $]0.99, 0.975[$

Teorema 8 *Considere las poblaciones dadas por la variables aleatorias* X, Y *que siguen una distribución normal con variancias poblacionales de* σ_1^2 *y* σ_2^2 *repectivamente. Sean* S_1^2 *y* S_2^2 *variancias muestras de muestas aleatorias independientes de tamaño* n_1 *y* n_2, *tomadas de cada población respectivamente. Se tiene que*

$$F = \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2}$$

sigue una distribución F *con* $v_1 = n_1 - 1$ *y* $v_2 = n_2 - 1$ *grados de libertad.*

6 Ejercicios

1. Utilice el app Probability Distribucions para determinar los siguientes valores

   (a) $\chi^2_{0.025,24}$ *R/* 12.4012

   (b) $t_{0.2,44}$ *R/* -0.849867

   (c) $t_{0.98,6}$ *R/* 2.61224

   (d) $f_{0.975,4,6}$ *R/* 6.227

   (e) $\chi^2_{0.2,22}$ *R/* 16.314

   (f) $z_{0.86}$ *R/* 1.08

   (g) $t_{0.1,30}$ *R/* -1.31042

   (h) $f_{0.05,12,22}$ *R/* 0.396354

2. Para cada una de las variables indicadas, determine la probabilidad solicitada utilizando el app Probability Distribucions y verifique que el valor es aproximadamente el dado en cada respuesta.

(a) $t \sim t(12), \quad P(t > 2)$ $R/ \quad 0.03$

(b) $t \sim t(14), \quad P(t < 1.7)$ $R/ \quad 0.95$

(c) $t \sim t(16), \quad P(t > -1.3)$ $R/ \quad 0.9$

(d) $t \sim t(18), \quad P(t < -1.9)$ $R/ \quad 0.04$

(e) $\chi^2 \sim \chi^2(20), \quad P(\chi^2 > 15)$ $R/ \quad 0.8$

(f) $\chi^2 \sim \chi^2(22), \quad P(\chi^2 < 25)$ $R/ \quad 0.8$

(g) $\chi^2 \sim \chi^2(24), \quad P(\chi^2 > 40)$ $R/ \quad 0.025$

3. Para cada una de las variables indicadas, determine la probabilidad solicitada utilizando el app Probability Distribucions y verifique que el valor encontrado se encuentra en el intervalo dado en la respuesta.

(a) $F \sim f(48, 20), \quad P(F > 2)$ $R/ \quad]0.025, 0.05[$

(b) $F \sim f(2, 10), \quad P(F > 0.03)$ $R/ \quad]0.95, 0.975[$

(c) $t \sim t(12), \quad P(t > 1.8)$ $R/ \quad]0.04, 0.05[$

(d) $\chi^2 \sim \chi^2(20), \quad P(\chi^2 < 19)$ $R/ \quad]0.2, 0.8[$

(e) $F \sim f(14, 20), \quad P(F > 0.57)$ $R/ \quad]0.1, 0.9[$

(f) $t \sim t(10), \quad P(t < -2)$ $R/ \quad]0.025, 0.04[$

(g) $\chi^2 \sim \chi^2(20), \quad P(\chi^2 > 10)$ $R/ \quad]0.95, 0.975[$

(h) $F \sim f(12, 18), P(F > 0.7)$ $R/ \quad]0.1, 0.9[$