

DISTRIBUCIONES MUESTRALES

Nexus-Probability

CURSO 2 (PROBABILIDAD II)

PARTE 5 / LECCIÓN 1

En esta lección, estudiaremos algunas distribuciones de probabilidad que surgen en la estadística y otras áreas donde aplicamos la probabilidad.

Recordemos la siguiente definición

Definición 1 (Muestra aleatoria) Una muestra aleatoria es una colección de variables aleatorias X_1, X_2, \dots, X_n , que cumplen con la condición de ser independientes y de tener cada una de ellas la misma distribución. Donde n es el tamaño de la muestra aleatoria.

Definición 2 Una estadística es una variable aleatoria de la forma $g(X_1, \dots, X_n)$ en donde X_1, \dots, X_n es una muestra aleatoria, y $g : \mathbb{R}^n \rightarrow \mathbb{R}$ es una función de Borel medible.

La **media muestral** es una estadística denotada por \bar{X} y definida como sigue

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

La **varianza muestral** es otro ejemplo importante de estadística, denotada por S^2 y definida como sigue

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

Ambos tienen la característica de ser estimadores insesgados de una distribución cualquiera. En particular, cuando la muestra aleatoria proviene de una distribución normal, resulta que la media y la varianza muestrales son independientes.

A continuación, estudiamos algunas distribuciones que surgen en la estadística al considerar funciones de una muestra aleatoria, en particular, la media y varianza muestral.

Distribución Ji-Cuadrada

La variable aleatoria continua X tiene una distribución ji-cuadrada con $n > 0$ grados de libertad, si su función de densidad es

$$f(x) = \begin{cases} \frac{1}{\Gamma(n/2)} \left(\frac{1}{2}\right)^{n/2} x^{n/2-1} e^{-x/2}, & \text{si } x > 0, \\ 0, & \text{si } x \leq 0. \end{cases}$$

En este caso, se escribe $X \sim \chi^2(n)$. El término χ^2 se lee ji-cuadrada. Es posible demostrar que $E(X) = n$, $Var(X) = 2n$. Observemos que la distribución $\chi^2(n)$ con $n = 2$ se reduce a la distribución $exp(\lambda)$ con $\lambda = 1/2$.

La suma de dos o mas variables aleatorias independientes con distribución ji-cuadrada es nuevamente una variable aleatoria ji-cuadrada, y sus grados de libertad son la suma de los grados de libertad de cada uno de los sumandos.

Definición 3 Sea X_1, \dots, X_m independientes tales que cada X_i tiene distribución $\chi^2(n_i)$, para $i = 1, \dots, m$. Entonces

$$\sum_{i=1}^m X_i \sim \chi^2(n_1 + \dots + n_m)$$

Definición 4 Sea X_1, \dots, X_n independientes cada una con distribución $N(\mu, \sigma^2)$. Entonces

$$\sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$$

Distribución t

La variable aleatoria continua X tiene una distribución t de Student con $n > 0$ grados de libertad si su función de densidad está dada por

$$f(x) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)} (1 + x^2/n)^{-(n+1)/2}$$

Se puede demostrar que $E(X) = 0$ y $Var(X) = n/(n-2)$ para $n > 2$. La primera igualdad nos dice que la distribución se encuentra siempre centrada en cero para cualquier valor del parámetro n .

Definición 5 Sean $X \sim N(0, 1)$ y $Y \sim \chi^2(n)$ independientes. Entonces

$$\frac{X}{\sqrt{Y/n}} \sim t(n)$$

El siguiente resultado es usado en estadística para efectuar estimaciones de la media de una población normal cuando la varianza es desconocida.

Definición 6 Sea X_1, \dots, X_n una m -a de una distribución $N(\mu, \sigma^2)$. Entonces

$$\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

Distribución F

La variable aleatoria continua X tiene una distribución F de Snedecor con parámetros $n > 0$ y $m > 0$ si su función de densidad es

$$f(x) = \begin{cases} \frac{\Gamma((n+m)/2)}{\Gamma(n/2)\Gamma(m/2)} \left(\frac{n}{m}\right)^{n/2} x^{n/2-1} \left(1 + \frac{n}{m}x\right)^{-(n+m)/2}, & \text{si } x > 0, \\ 0, & \text{si } x \leq 0. \end{cases}$$

Se escribe $X \sim F(n, m)$. Puede demostrarse,

$$E(X) = \frac{m}{m-2} \quad \text{para } m > 2$$

$$Var(X) = \frac{2m^2(m+n-2)}{n(m-2)^2(m-4)} \quad \text{para } m > 4$$

Los siguientes dos resultados indican la forma en la que se obtiene esta distribución.

Definición 7 Sean $X \sim \chi^2(n)$ y $Y \sim \chi^2(m)$ independientes. Entonces,

$$\frac{X/n}{Y/m} \sim F(n, m)$$

Definición 8 Si $X \sim t(n)$, entonces $X^2 \sim F(1, n)$

Ejercicios

Los siguientes ejercicios propuestos son algunas definiciones mencionadas anteriormente.

Ejercicio 1 Sea X_1, \dots, X_m independientes tales que cada X_i tiene distribución $\chi^2(n_i)$, para $i = 1, \dots, m$. Entonces

$$\sum_{i=1}^m X_i \sim \chi^2(n_1 + \dots + n_m)$$

Solución.

Demostración del resultado anterior.

Es suficiente demostrar el resultado para el caso de dos variables aleatorias. Sean X y Y independientes con distribución ji-cuadrada con grados de libertad n y m , respectivamente. Este ligero cambio en la notación evitará el uso de subíndices.

$$\begin{aligned} f_{X+Y} &= \int_0^u f_X(u-v)f_Y(v)dv \\ &= \int_0^u \frac{1}{\Gamma(n/2)}(1/2)^{n/2}(u-v)^{n/2-1}e^{-(u-v)/2} \\ &= \frac{1}{\Gamma(m/2)}(1/2)^{m/2}v^{m/2-1}e^{-v/2}dv \\ &= \frac{1}{\Gamma(n/2)\Gamma(m/2)}(1/2)^{(n+m)/2}e^{-u/2} \int_0^u (u-v)^{n/2-1}v^{m/2-1}dv \end{aligned}$$

Haciendo el cambio de variable $w(v) = v/u$ se obtiene

$$f_{X+Y}(u) = \frac{1}{\Gamma(n/2)\Gamma(m/2)}(1/2)^{(n+m)/2}e^{-u/2}u^{(n+m)/2-1} \int_0^u (1-w)^{n/2-1}w^{m/2-1}dw$$

La integral resultante es $B(n/2, m/2)$. Entonces

$$\begin{aligned} f_{X+Y}(u) &= \frac{B(n/2, m/2)}{\Gamma(n/2)\Gamma(m/2)}(1/2)^{(n+m)/2}e^{-u/2}u^{(n+m)/2-1} \\ &= \frac{1}{\Gamma((n+m)/2)}(1/2)^{(n+m)/2}e^{-u/2}u^{(n+m)/2-1} \end{aligned}$$

Esta es la función de densidad de la distribución $\chi^2(n+m)$

Ejercicio 2 Sea X_1, \dots, X_n independientes cada una con distribución $N(\mu, \sigma^2)$. Entonces

$$\sum_{i=1}^n \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi^2(n)$$

Solución.

Esto es una consecuencia sencilla de las dos proposiciones anteriores. Como cada una de las variables X_i tiene distribución $N(\mu, \sigma^2)$, para $i = 1, \dots, n$, entonces $(X_i - \mu)/\sigma$ tiene distribución $N(0, 1)$. Por lo tanto $(X_i - \mu)^2/\sigma^2$ tiene distribución $\chi^2(1)$. En consecuencia $\sum_{i=1}^n (X_i - \mu)^2/\sigma^2$ tiene distribución $\chi^2(n)$.