

Capítulo 7 Distribuciones muestrales y el teorema del límite central

7.1 Introducción

A lo largo de este capítulo trabajaremos con funciones de las variables Y_1, Y_2, \dots, Y_n , observadas en una muestra aleatoria seleccionada de una población de interés. Estas variables son independientes y tienen la misma distribución. Se pueden utilizar funciones de las variables aleatorias observadas en una muestra para calcular o tomar decisiones acerca de parámetros desconocidos de la población. Como se vio en el capítulo anterior, si se crea una función de variables aleatorias entonces esta función es una variable aleatoria. Con esto se puede llegar a la siguiente definición.

Definición 7.1

Un *estadístico* es una función de las variables aleatorias observadas en una muestra y constantes conocidas.

Algunos ejemplos de estadísticos son: la media muestral \bar{Y} , la varianza muestral S^2 o la amplitud $R = Y_{(n)} - Y_{(1)}$. Se utilizan estadísticos para hacer inferencias acerca de parámetros de una población. Además como son funciones de variables aleatorias entonces los estadísticos también son variables aleatorias. En consecuencia, todos los estadísticos tienen distribuciones de probabilidad que llamaremos *distribuciones muestrales*. Esta distribución dependerá de la distribución de las variables aleatorias observables de la muestra.

7.2 Distribuciones muestrales relacionadas con la distribución normal

Se ha visto anteriormente que muchos fenómenos observados en el mundo real tienen distribuciones de frecuencia relativas que se pueden modelar usando la distribución de probabilidad normal. Entonces supongamos que las variables aleatorias observables en una muestra aleatorias son Y_1, Y_2, \dots, Y_n y que sean independientes con función de densidad normal. Entonces se puede encontrar un estadístico que ayude a hacer inferencias acerca de la población. Esto se observa en el siguiente teorema.

Teorema 7.1

Sea Y_1, Y_2, \dots, Y_n una muestra aleatorias de tamaño n de una distribución normal con media μ y varianza σ^2 . Entonces

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$$

está distribuida normalmente con media $\mu_{\bar{Y}} = \mu$ y varianza $\sigma_{\bar{Y}}^2 = \sigma^2/n$.

Demostración

Como Y_1, Y_2, \dots, Y_n es una muestra aleatoria de una distribución normal con media μ y varianza σ^2 , Y_i , $i = 1, 2, \dots, n$ son variables aleatorias independientes distribuidas normalmente, con $E(Y_i) = \mu$ y $V(Y_i) = \sigma^2$. Además,

$$\begin{aligned} \bar{Y} &= \frac{1}{n} \sum_{i=1}^n Y_i = \frac{1}{n} (Y_1) + \frac{1}{n} (Y_2) + \dots + \frac{1}{n} (Y_n) \\ &= a_1 Y_1 + a_2 Y_2 + \dots + a_n Y_n \end{aligned}$$

donde $a_i = 1/n$, $i = 1, 2, \dots, n$. Así, \bar{Y} es una combinación lineal de Y_1, Y_2, \dots, Y_n , y se puede aplicar el Teorema 6.3 para concluir que \bar{Y} está distribuida normalmente con

$$\begin{aligned} E(\bar{Y}) &= E \left[\frac{1}{n} (Y_1) + \dots + \frac{1}{n} (Y_n) \right] \\ &= \frac{1}{n} \mu + \dots + \frac{1}{n} \mu \\ &= \mu \end{aligned}$$

y

$$\begin{aligned} V(\bar{Y}) &= V \left[\frac{1}{n} (Y_1) + \dots + \frac{1}{n} (Y_n) \right] \\ &= \frac{1}{n^2} (\sigma^2) + \dots + \frac{1}{n^2} (\sigma^2) \\ &= \frac{1}{n^2} (n\sigma^2) = \frac{\sigma^2}{n} \end{aligned}$$

Esto es, la distribución muestral de \bar{Y} es normal con media $\mu_{\bar{Y}} = \mu$ y varianza $\sigma_{\bar{Y}}^2 = \sigma^2/n$.

De acuerdo con estas condiciones si \bar{Y} está normalmente distribuida con media $\mu_{\bar{Y}} = \mu$ y varianza $\sigma_{\bar{Y}}^2 = \sigma^2/n$. Se deduce que,

$$Z = \frac{\bar{Y} - \mu_{\bar{Y}}}{\sigma_{\bar{Y}}} = \frac{\bar{Y} - \mu}{\sigma/\sqrt{n}} = \sqrt{n} \left(\frac{\bar{Y} - \mu}{\sigma} \right)$$

tiene una distribución normal estándar. Con esto se puede llegar al siguiente teorema.

Teorema 7.2

Si Y_1, Y_2, \dots, Y_n está definida como en el Teorema 7.1. Entonces $Z_i = (Y_i - \mu)/\sigma$ son variables aleatorias independientes con distribución normales estándar, $i = 1, 2, \dots, n$, y

$$\sum_{i=1}^n Z_i^2 = \sum_{i=1}^n \left(\frac{Y_i - \mu}{\sigma} \right)^2$$

tienen una distribución χ^2 con n grados de libertad (gl).

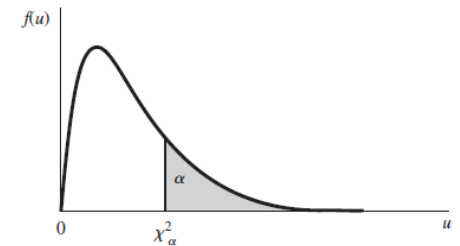
Demostración

Como Y_1, Y_2, \dots, Y_n es una muestra aleatoria de una distribución normal con media μ y varianza σ^2 . Como ya se ha visto anteriormente $Z_i = (Y_i - \mu)/\sigma$ tiene una distribución normal estándar para $i = 1, 2, \dots, n$. Además las variables aleatorias Z_i son independientes porque las variables aleatorias Y_i son independiente, $i = 1, 2, \dots, n$. El hecho de que $\sum_{i=1}^n Z_i^2$ tiene una distribución χ^2 con n grados de libertad se deduce directamente del Teorema 6.4

La distribución χ^2 es bastante importante en la estadística por la relación que tiene con la distribución normal. Debido a su importancia se tienen tablas de valores para esta distribución para distintos niveles de confianza y grados de libertad. En general se tiene,

$$P(\chi^2 > \chi_{\alpha}^2) = \alpha \quad P(\chi^2 \leq \chi_{\alpha}^2) = 1 - \alpha$$

y que $\chi_{\alpha}^2 = \phi_{1-\alpha}$, el cuantil $(1 - \alpha)$ de la variable aleatoria χ^2 . El área sombreada de la siguiente figura muestra esta probabilidad.



Otro ejemplo del papel importante de la distribución χ^2 es el siguiente. Se demostrará en el capítulo 8 que un buen estimador de σ^2 es la varianza muestral,

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

El siguiente teorema proporciona la distribución de probabilidad para una función del estadístico S^2 .

Teorema 7.3

Sea Y_1, Y_2, \dots, Y_n una muestra aleatoria de una distribución normal con media μ y varianza σ^2 . Entonces

$$\frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^n (Y_i - \bar{Y})^2$$

tiene una distribución χ^2 con $(n-1)$ grados de libertad. También \bar{Y} y S^2 son variables independientes.

Demostración

La demostración completa de este teorema aparece en el ejercicio 13.93. Para entender mejor el resultado general, consideraremos el caso $n = 2$ y demostraremos que $(n-1)S^2/\sigma^2$ tiene una distribución χ^2 con 1 grado de libertad. En el caso $n = 2$,

$$\bar{Y} = (1/2)(Y_1 + Y_2),$$

y por lo tanto,

$$\begin{aligned} S^2 &= \frac{1}{2-1} \sum_{i=1}^2 (Y_i - \bar{Y})^2 \\ &= \left[Y_1 - \frac{1}{2}(Y_1 + Y_2) \right]^2 + \left[Y_2 - \frac{1}{2}(Y_1 + Y_2) \right]^2 \\ &= \left[\frac{1}{2}(Y_1 - Y_2) \right]^2 + \left[\frac{1}{2}(Y_2 - Y_1) \right]^2 \\ &= 2 \left[\frac{1}{2}(Y_1 - Y_2) \right]^2 \\ &= \frac{(Y_1 - Y_2)^2}{2} \end{aligned}$$

Se deduce que, cuando $n = 2$,

$$\frac{(n-1)S^2}{\sigma^2} = \frac{(Y_1 - Y_2)^2}{2\sigma^2} = \left(\frac{Y_1 - Y_2}{\sqrt{2}\sigma} \right)^2$$

Demostraremos que esta cantidad es igual al cuadrado de una variable aleatoria normal estándar; es decir, que se trata de una variable Z^2 que, como ya hemos demostrado en que posee una distribución χ^2 con 1 grado de libertad. Como $Y_1 - Y_2$ es una combinación de variables aleatorias independientes distribuidas normalmente ($Y_1 - Y_2 = a_1 Y_1 + a_2 Y_2$ con $a_1 = 1$ y $a_2 = -1$), el Teorema 6.3 nos dice que $Y_1 - Y_2$ tiene una distribución normal con media $1\mu - 1\mu = 0$ y varianza $(1)^2\sigma^2 + (-1)^2\sigma^2 = 2\sigma^2$. Por lo tanto,

Demostración

$$Z = \frac{Y_1 - Y_2}{\sqrt{2}\sigma}$$

tiene una distribución normal estándar. Como para $n = 2$

$$\frac{(n-1)S^2}{\sigma^2} = \left(\frac{Y_1 - Y_2}{\sqrt{2}\sigma} \right)^2 = Z^2$$

se deduce que $(n-1)S^2/\sigma^2$ tiene una distribución χ^2 con 1 grado de libertad.

En el Ejemplo 6.13 demostramos que $U_1 = (Y_1 + Y_2)/\sigma$ y $U_2 = (Y_1 - Y_2)/\sigma$ son variables aleatorias independientes. Observe que, debido a que $n = 2$,

$$\bar{Y} = \frac{Y_1 + Y_2}{2} = \frac{\sigma U_1}{2} \quad S^2 = \frac{(Y_1 - Y_2)^2}{2} = \frac{(\sigma U_2)^2}{2}$$

Como \bar{Y} solo es una función de U_1 y S^2 es una función de U_2 , la independencia de U_1 y U_2 implica la independencia de \bar{Y} y S^2 .

El resultado del Teorema 7.1 proporciona la base para el desarrollo de procedimientos que permiten hacer inferencias acerca de la media μ de una población normal con varianza conocida σ^2 . En dicho caso el Teorema 7.1 indica que $\frac{\sqrt{n}(\bar{Y} - \mu)}{\sigma}$ tiene una distribución normal estándar. Cuando σ no se conoce, puede ser estimada con $S = \sqrt{S^2}$ y la cantidad

$$\sqrt{n} \left(\frac{\bar{Y} - \mu}{S} \right)$$

proporciona la base para desarrollar métodos de inferencia respecto de μ . Demostraremos que esta expresión tiene una distribución conocida como *distribución t de Student* con $n-1$ grados de libertad. Esta se define de la siguiente manera.

Definición 7.2

Sea Z una variable aleatoria normal estándar y sea W una variable con distribución χ^2 con v grados de libertad. Entonces, si W y Z son independientes,

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

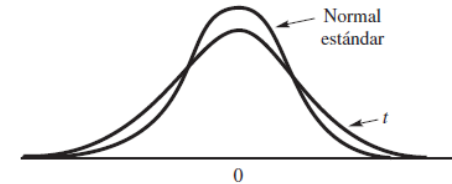
se dice que tiene una *distribución t* con v grados de libertad.

Si Y_1, Y_2, \dots, Y_n constituye una muestra aleatoria de una población normal con media μ y varianza σ^2 , el Teorema 7.1 puede aplicarse para demostrar que $Z = \sqrt{n}(\bar{Y} - \mu)/\sigma$ tiene una distribución normal estándar. El teorema 7.3 nos dice que $W = (n-1)S^2/\sigma^2$ tiene una distribución χ^2 con $v = n-1$ grados de libertad y que Z y W

son independientes (puesto que \bar{Y} y S^2 son independientes). Entonces con la Definición 7.2, se tiene lo siguiente,

$$\begin{aligned} T &= \frac{Z}{\sqrt{W/v}} = \frac{\sqrt{n}(\bar{Y} - \mu)/\sigma}{\sqrt{[(n-1)S^2/\sigma^2]/(n-1)}} \\ &= \frac{\sqrt{n}(\bar{Y} - \mu)/\sigma}{\sqrt{S^2/\sigma^2}} = \sqrt{n} \left(\frac{\bar{Y} - \mu}{S} \right) \end{aligned}$$

tiene una distribución t con $(n-1)$ grados de libertad. Al igual que la función de densidad normal estándar, la función de densidad t es simétrica alrededor de cero. Además para $v > 1$, $E(T) = 0$, y para $v > 2$, $V(T) = v/(v-2)$. De esta manera, vemos que, si $v > 1$, una variable aleatoria con distribución t tiene el mismo valor esperado que una variable aleatoria normal estándar. No obstante, una variable normal estándar siempre tiene una varianza 1 mientras que, si $v > 2$, la varianza de una variable aleatoria con una distribución t siempre es mayor que 1. En la siguiente figura se muestran ambas distribuciones.



Ahora veremos otra aplicación de la distribución χ^2 , suponga que queremos comparar las varianzas de dos poblaciones normales con base en información contenida en muestras aleatorias independientes provenientes de las dos poblaciones. Suponga que se toman muestras de tamaño n_1 y n_2 de las dos poblaciones con varianzas σ_1^2 y σ_2^2 , respectivamente. Si calculamos la varianza muestral para la primera muestra entonces S_1^2 calcula σ_1^2 . Del mismo modo para la segunda muestra se tiene que S_2^2 calcula σ_2^2 . Al dividir ambas varianzas muestrales entonces esta razón podría usarse para hacer inferencias acerca de las magnitudes relativas de σ_1^2 y σ_2^2 . Si dividimos cada S_i^2 entre σ_i^2 entonces la razón resultante

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \left(\frac{\sigma_2^2}{\sigma_1^2} \right) \left(\frac{S_1^2}{S_2^2} \right)$$

tiene una *distribución F* con (n_1-1) grados de libertad en el numerador y (n_2-1) grados de libertad en el denominador. La definición de esta variable es la siguiente.

Definición 7.3

Sean W_1 y W_2 variables aleatorias independientes con distribución χ^2 , con v_1 y v_2 grados de libertad, respectivamente. Entonces se dice que

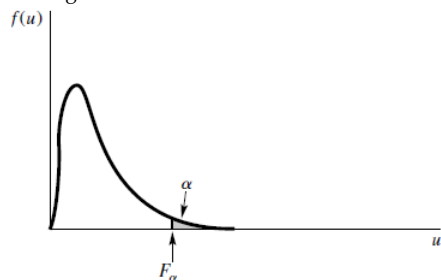
$$F = \frac{W_1/v_1}{W_2/v_2}$$

tiene una *distribución F* con v_1 grados de libertad en el numerador y v_2 grados de libertad en el denominador.

Consideremos dos muestras aleatorias normales tomadas de distribuciones normales, sabemos que $W_1 = (n_1 - 1)S_1^2/\sigma_1^2$ y $W_2 = (n_2 - 1)S_2^2/\sigma_2^2$ tienen distribuciones χ^2 independientes con $v_1 = (n_1 - 1)$ y $v_2 = (n_2 - 1)$ grados de libertad, respectivamente. Entonces la definición anterior implica lo siguiente.

$$F = \frac{W_1/v_1}{W_2/v_2} = \frac{[(n_1 - 1)S_1^2/\sigma_1^2]/(n_1 - 1)}{[(n_2 - 1)S_2^2/\sigma_2^2]/(n_2 - 1)} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$$

tiene una distribución F con $(n_1 - 1)$ grados de libertad en el numerador y $(n_2 - 1)$ grados de libertad en el denominador. Se puede demostrar que esta distribución tiene $E(F) = v_2/(v_2 - 2)$ si $v_2 > 2$. También si $v_2 > 4$, entonces $V(F) = [2v_2^2(v_1 + v_2 - 2)]/[v_1(v_2 - 2)^2(v_2 - 4)]$. Observamos que la media de una variable aleatorias con distribución F depende sólo del número de grados de libertad v_2 del denominador. Se observa una distribución F en la siguiente figura.



7.3 Teorema del límite central

Un resultado importante es de la estadística es el *Teorema del límite central*, este resultado nos ayuda cuando hay variables aleatorias que no tienen una distribución normal.

Teorema 7.4

Teorema del límite central: Sean Y_1, Y_2, \dots, Y_n variables aleatorias independientes y distribuidas idénticamente con $E(Y_i) = \mu$ y $V(Y_i) = \sigma^2 < \infty$. Definamos

$$U_n = \frac{\sum_{i=1}^n Y_i - n\mu}{\sigma\sqrt{n}} = \frac{\bar{Y} - \mu}{\sigma/\sqrt{n}}$$

donde $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$. Entonces la función de distribución de U_n converge hacia la función de distribución normal estándar cuando $n \rightarrow \infty$. Esto es,

$$\lim_{n \rightarrow \infty} P(U_n \leq u) = \int_{-\infty}^u \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

para toda u .

En otras palabras este teorema implica que conforme incrementa en tamaño una muestra aleatoria, esta empieza a tener propiedades

similares a la distribución normal. Esto se puede expresar de otra manera más simple. Sea el estadístico \bar{Y} , se puede decir que esta *distribuido normalmente en forma asintótica* con media μ y varianza σ^2/n . El teorema del límite central se puede aplicar a una muestra aleatoria Y_1, Y_2, \dots, Y_n para *cualquier* distribución mientras que $E(Y_i) = \mu$ y $V(Y_i) = \sigma^2$ sean finitas y el tamaño de muestra sea grande.

7.4 Una demostración del teorema del límite central

Teorema 7.5

Sean Y y Y_1, Y_2, Y_3, \dots , variables aleatorias con funciones generadoras de momento $m(t)$ y $m_1(t), m_2(t), m_3(t), \dots$, respectivamente. Si

$$\lim_{n \rightarrow \infty} m_n(t) = m(t)$$

para toda t real, entonces la función de distribución de Y_n converge hacia la función de distribución de Y cuando $n \rightarrow \infty$.

A continuación se dará la demostración del teorema del límite central, Teorema 7.4.

Demostración

Sea

$$U_n = \sqrt{n} \left(\frac{\bar{Y} - \mu}{\sigma} \right) = \frac{1}{\sqrt{n}} \left(\frac{\sum_{i=1}^n Y_i - n\mu}{\sigma} \right)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i$$

donde $Z_i = \frac{Y_i - \mu}{\sigma}$. Debido a que las Y_i de las variables aleatorias son independientes e idénticamente distribuidas, $Z_i, i = 1, 2, \dots, n$, son independientes e idénticamente distribuidas con $E(Z_i) = 0$ y $V(Z_i) = 1$. Como la función generadora de momento de la suma de variables aleatorias independientes es el producto de las funciones generadoras de momentos de cada una,

$$m_{\sum Z_i}(t) = m_{Z_1}(t) \times m_{Z_2}(t) \times \dots \times m_{Z_n}(t) = [m_{Z_1}(t)]^n$$

y

$$m_{U_n}(t) = m_{\sum Z_i} \left(\frac{t}{\sqrt{n}} \right) = \left[m_{Z_1} \left(\frac{t}{\sqrt{n}} \right) \right]^n$$

Demostración

Por el teorema de Taylor, con residuo

$$m_{Z_1}(t) = m_{Z_1}(0) + m'_{Z_1}(0)t + m''_{Z_1}(\xi) \frac{t^2}{2},$$

donde $0 < \xi < t$ y como $m_{Z_1}(0) = E(e^{0Z_1}) = E(1) = 1$ y $m'_{Z_1}(0) = E(Z_1) = 0$,

$$m_{Z_1}(t) = 1 + \frac{m''_{Z_1}(\xi)}{2} t^2$$

donde $0 < \xi < 1$. Por lo tanto,

$$\begin{aligned} m_{U_n}(t) &= \left[1 + \frac{m''_{Z_1}(\xi_n)}{2} \left(\frac{t}{\sqrt{n}} \right)^2 \right]^n \\ &= \left[1 + \frac{m''_{Z_1}(\xi_n) t^2 / 2}{n} \right]^n \end{aligned}$$

donde $0 < \xi_n < \frac{t}{\sqrt{n}}$. Observe que conforme $n \rightarrow \infty$, $\xi_n \rightarrow 0$ y $m''_{Z_1}(\xi_n) t^2 / 2 \rightarrow m''_{Z_1}(0) t^2 / 2 = E(Z_1^2) t^2 / 2 = t^2 / 2$ por que $E(Z_1^2) = V(Z_1) = 1$. Recuerde que si

$$\lim_{n \rightarrow \infty} b_n = b \quad \text{entonces} \quad \lim_{n \rightarrow \infty} \left(1 + \frac{b_n}{n} \right)^n = e^b$$

Finalmente,

$$\lim_{n \rightarrow \infty} m_{U_n}(t) = \lim_{n \rightarrow \infty} \left[1 + \frac{m''_{Z_1}(\xi_n) t^2 / 2}{n} \right]^n = e^{t^2/2},$$

es la función generadora de momento para una variable aleatoria normal estándar. Con la aplicación del Teorema 7.5 concluimos que U_n tiene una función de distribución que converge hacia la función de distribución de la variable aleatoria normal estándar.

7.5 Aproximación normal a la distribución binomial

El teorema del límite central también se puede usar para aproximar probabilidades de algunas variables aleatorias discretas cuando las probabilidades exactas sean difíciles de calcular. Por ejemplo, suponemos que Y tiene una distribución binomial con n intentos y probabilidad de éxito en cualquier intento denotado por p . Si n es demasiado grande encontrar la probabilidad de algún valor cercano a n se vuelve una tarea muy difícil. De manera alternativa, podemos ver a Y , el número de éxitos en n intentos, como una suma de una muestra formada de ceros y unos, esto es,

$$Y = \sum_{i=1}^n X_i$$

donde

$$X_i = \begin{cases} 1, & \text{si el } i\text{-ésimo intento resulta en éxito,} \\ 0, & \text{en otro caso} \end{cases}$$

Las variables aleatorias X_i para $i = 1, 2, \dots, n$ son independientes (porque los intentos son independientes) y es fácil demostrar que $E(X_i) = p$ y $V(X_i) = p(1-p)$ para $i = 1, 2, \dots, n$. En consecuencia, cuando n es grande, la fracción muestral de éxitos,

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

posee una distribución muestral aproximadamente normal con media $E(X_i) = p$ y varianza $V(X_i)/n = p(1-p)/n$.

Entonces, hemos empleado el Teorema 7.4 para establecer que si Y es una variable aleatoria binomial con parámetros n y p si n es grande, entonces Y/n tiene aproximadamente la misma distribución que U , donde U está distribuida normalmente con media $\mu_U = p$ y varianza $\sigma_U^2 = p(1-p)/n$. De la misma manera para n grande, podemos considerar que Y tiene aproximadamente la misma distribución que W , donde W está distribuida normalmente con media $\mu_W = np$ y varianza $\sigma_W^2 = np(1-p)$.

Sin embargo el aplicar este teorema no siempre funciona bien. Para que funcione n necesita ser relativamente grande y que p no sea cer-

cano a cero o uno. Una regla práctica útil es que la aproximación normal a la distribución binomial es apropiada cuando $p \pm 3\sqrt{pq/n}$ está en el intercalo $(0,1)$, es decir, si

$$0 < p - 3\sqrt{pq/n} \quad \text{y} \quad p + 3\sqrt{pq/n} < 1$$

Además se puede demostrar un criterio más conveniente pero similar que dice que la aproximación normal es adecuada si

$$n > 9 \left(\frac{\text{el mayor de } p \text{ y } q}{\text{el menor de } p \text{ y } q} \right)$$