Unidad 8

- Distribución muestral.

Capítulo 7

DISTRIBUCIÓN MUESTRAL

7.1 INTRODUCCIÓN

Como es de esperarse, a partir de una misma población se pueden tomar muchas muestras diferentes del mismo tamaño.

El desarrollo del ejercicio que sigue nos indicará cómo se obtienen los valores muestrales de una variable (la media muestral en este caso) de acuerdo con los datos particulares observados. Además, nos enseñará cómo se obtienen los valores de probabilidad para cada uno de los valores muestrales, con lo cual se llega así a lo que se conoce como *distribución muestral*.

Suponga que la variable aleatoria X toma cualquiera de los cuatro valores 2, 4, 6, 8. Estos cuatro números se pueden utilizar para simular una población infinita si se toman muestras ordenadas con repetición.

Suponga además que se toman de esta población muestras de tamaño 2. Se puede escribir los números 2, 4, 6 y 8 en cuatro bolas y valerse del procedimiento de la bolsa para la selección de las muestras. Una muestra puede consistir en los números 2 y 8, y otra puede estar formada por los números 8 y 6, o 6 y 2 o alguna otra combinación. Es de gran importancia conocer todas las muestras diferentes que podrían resultar del experimento.

En total existen 16 muestras posibles que se pueden seleccionar de esta población. La tabla 7.1 arroja todas las 16 muestras y en ella X_1 designa el número de la primera bola que se saca, X_2 el número

de la segunda bola. Si bien podría obtenerse hipotéticamente un número infinito de muestras, solamente podrían resultar 16 muestras posibles diferentes. Es claro que si se toman miles de tales muestras, muchas de ellas serían idénticas.

En la tabla 7.1 también se da la media aritmética de cada muestra. La media muestral $\overline{X}$ toma cualquiera de los valores del conjunto de siete valores 2, 3, 4, 5, 6, 7, 8. Estos se designan con $\overline{x}$; son todos los valores posibles que puede tomar $\overline{X}$. De las 16 muestras posibles, una tiene media 2; dos tienen media 3; tres media 4; cuatro media 5; tres media 6; dos media 7 y una media 8.

Tabla 7.1 *Diferentes muestras posibles de tamaño 2 de X distribuidas uniformemente sobre {2,4,6,8}.*

Muestra	X_1	X_2	*Media muestral* $\overline{X}$
1	2	2	2
2	2	4	3
3	2	6	4
4	2	8	5
5	4	2	3
6	4	4	4
7	4	6	5
8	4	8	6
9	6	2	4
10	6	4	5
11	6	6	6
12	6	8	7
13	8	2	5
14	8	4	6
15	8	6	7
16	8	8	8

A partir de la tabla 7.1 obtenemos la tabla 7.2 en donde aparecen los valores de $\overline{X}$ junto a sus respectivas probabilidades; esto es, la distribución de $\overline{X}$. Para la obtención de estas probabilidades, se debe tener presente que al efectuarse un muestreo sin repetición cada elemento de la muestra tiene una probabilidad de $\frac{1}{4}$ de ser escogido y así, cada muestra de tamaño dos tiene probabilidad de $\left(\frac{1}{4}\right)\left(\frac{1}{4}\right) = \frac{1}{16}$ de darse.

De acuerdo con la tabla 7.2, la probabilidad de obtener una media muestral 2 es $\frac{1}{16}$, la de 3 es $\frac{2}{16}$, y así sucesivamente. Pero

Tabla 7.2 *Distribución de la media muestral $\overline{X}$ con tamaño de muestra 2.*

Media muestral $\overline{x}$	Número de muestras	Probabilidad $P(\overline{x})$
2	1	$\frac{1}{16}$
3	2	$\frac{2}{16}$
4	3	$\frac{3}{16}$
5	4	$\frac{4}{16}$
6	3	$\frac{3}{16}$
7	2	$\frac{2}{16}$
8	1	$\frac{1}{16}$
Total	16	1.0

Fuente: Tabla 7.1

esto no quiere decir que de 1,600 muestras que se tomen, 100 tengan media 2 y 200 muestras tengan media 3, etc. Ni tampoco que de 16 millones de muestras se conseguiría exactamente un millón con media 1 y dos millones con media 3. Sin embargo, si se toman 16,000 millones de muestras, el número de ellas con media 1 se acercará a 1,000 millones o sea que la razón será aproximadamente $\frac{1}{16}$. Es decir, que la distribución de probabilidades es la de frecuencia relativa *teórica* y sólo se la puede aproximar mediante un número infinitamente grande de pruebas. (No hay para qué decir que la distribución de probabilidades aquí indicada es una ilustración de una distribución teórica. En la práctica jamás se toman tantas muestras, sino una especial, la cual es una parte de la distribución. Con base en la muestra se hacen inferencias acerca de las características de la distribución).

La distribución muestral de la media que se indica en la tabla 7.2 se representa gráficamente en la figura 7.1.

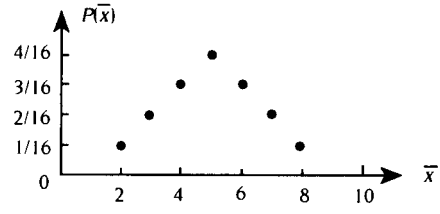


Figura 7.1 *Distribución de probabilidad de la media muestral de la tabla 7.2.*

El construir la distribución de la media muestral con enumeración de todas las diferentes muestras posibles, es naturalmente un proce-

dimiento inefectivo y a veces hasta imposible. Es inefectivo, puesto que habrían de enumerarse demasiadas muestras diferentes cuando la variable aleatoria tiene un espacio muestral grande.

Por ejemplo, si el tamaño de las muestras es 10 y se toman de una población de sólo 100 elementos distintos, habría $\binom{100}{10} = 17{,}310{,}309{,}456{,}440$ muestras diferentes no ordenadas. Es imposible, puesto que habría un número infinito de muestras diferentes si la variable aleatoria tiene un espacio muestral infinito. Es por tanto más conveniente obtener la distribución de una estadística muestral tal como la media, en función de las probabilidades teóricas de todos los resultados muestrales posibles cuando se hacen suposiciones sobre la población.

Por consiguiente, aun cuando se pueda construir una distribución de frecuencias relativa empírica de una estadística muestral al tomar un gran número de muestras de igual tamaño de la misma población y luego distribuirlas de acuerdo con los resultados muestrales efectivos, es más corriente establecer la distribución de una estadística muestral como una relación funcional entre todos los resultados posibles y sus probabilidades correspondientes.

A continuación nos proponemos precisar los conceptos que nos conducirán a establecer la distribución de una estadística sin necesidad de extraer muestras para determinar la frecuencia relativa, cuestión que, como ya se comentó, es inefectiva y hasta imposible.

7.2 DISTRIBUCIÓN CONJUNTA. INDEPENDENCIA DE VARIABLES

Recordemos que si se tiene una variable aleatoria discreta X, entonces la expresión o escritura $[X = x]$ representa un evento del espacio muestral S sobre el cual está definida la variable. Igualmente, si sobre el mismo espacio muestral S definimos otra variable aleatoria Y, entonces la escritura $[Y = y]$ representa también un evento de S. De lo anterior podemos señalar que una escritura tal como:
$[X = x, Y = y]$ representa la intersección del evento indicado por $[X = x]$ con el evento $[Y = y]$. Esto es, $[X = x, Y = y] = [X = x] \cap [Y = y]$ (7-1)

También se ha dicho (capítulo 4) que cuando se afirma que dos eventos A y B son independientes, entonces debe cumplirse que $P[A \cap B] = P[A]\,P[B]$. Ahora bien, trasladada esta condición a (7-1) tenemos $P[X = x, Y = y] = P[X = x]\,P[Y = y]$ y llegamos a la siguiente consideración:

Dadas dos variables aleatorias discretas X *y* Y, *decimos que son independientes si se cumple que*

$$P[X = x, Y = y] = P[X = x]\,P[Y = y] \tag{7-2}$$

La expresión que aparece al lado izquierdo de (7-2) recibe el nombre de *distribución conjunta de X y Y;* cada uno de los factores que aparecen a la derecha de (7-2) recibe el nombre de *distribución marginal de X y Y* respectivamente. Y dentro de esta

terminología nos expresamos al decir: *X y Y son independientes cuando la distribución conjunta es igual al producto de las distribuciones marginales.* En realidad, al inicio del capítulo, cuando se consideró el ejercicio introductorio, se hizo uso de la citada regla.

Intuitivamente, dos variables aleatorias X y Y se dicen independientes cuando los valores que asume una de ellas no influye ni está influenciada por los valores de la otra.

Entre dos variables aleatorias X y Y se pueden realizar las operaciones básicas de aritmética como son la adición (resta) y multiplicación (división) y tenemos: La variable suma se representa por $X + Y$ y la variable diferencia $X - Y$. De igual forma la variable producto se representa por XY y la variable cociente $\frac{X}{Y}$

El concepto de valor esperado puede extenderse para el caso de una suma de variables (de hecho ya se hizo en el capítulo 5), de una resta, de un producto y de un cociente y tenemos que:
"Si X es una variable aleatoria discreta que asume los valores $x_1, x_2, \ldots, x_n$; Y otra variable aleatoria que asume los valores $y_1, y_2, \ldots, y_m$", entonces:

$$\text{i) } E(X+Y) = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i + y_j)P[X = x_i,\ Y = y_j] \tag{7-3}$$

$$\text{ii) } E(XY) = \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i y_j)P[X = x_i, Y = y_j] \tag{7-4}$$

Nota. La escritura "doble sumatoria" $\Sigma\Sigma$ quiere decir que debemos considerar todos los productos posibles entre los valores de X y los valores de Y y luego sumar los productos de estos resultados con la probabilidad conjunta.

Llegados a este punto es conveniente que desarrollemos un ejercicio para ver con mayor claridad lo que hemos expuesto hasta ahora.

Supongamos que lanzamos un par de tetraedros[1] y consideremos las variables aleatorias X, Y definidas de la manera siguiente:

X = Número de puntos que muestra la cara sobre la cual se apoya el primer tetraedro.

Y = Número de puntos que muestra la cara sobre la cual se apoya el segundo tetraedro.

Los valores posibles de X son 1, 2, 3, 4; lo mismo que los de Y. Ahora bien, al considerar las dos variables conjuntamente obtenemos las parejas de valores 1 - 1, 1 - 2, 1 - 3, 1 - 4, ... , 4 - 1, 4 - 2, 4 - 3, 4 - 4 y así el espacio muestral está dado por $S = \{1-1, 1-2, 1-3, 1-4, \ldots, 4-1, 4-2, 4-3, 4-4\}$.

En el primer cuadro (en su parte interior) se registran los distintos resultados de la variable suma $X + Y$. En el segundo cuadro se representa la variable producto XY. Los números que aparecen en la primera columna y primera fila de ambos cuadros representan los valores posibles de X y de Y.

[1] Desde el punto de vista de la probabilidad no existe diferencia entre lanzar dos tetraedros una vez y lanzar un tetraedro dos veces.

+	1	2	3	4
1	2	3	4	5
2	3	4	5	6
3	4	5	6	7
4	5	6	7	8

Cuadro 1 *Valores de X + Y.*

×	1	2	3	4
1	1	2	3	4
2	2	4	6	8
3	3	6	9	12
4	4	8	12	16

Cuadro 2 *Valores de XY.*

Ahora vamos a obtener la distribución de probabilidad para la variable $X + Y$. Para ello observe en primer lugar que los valores distintos de $X + Y$ son: 2 que se da 1 vez; 3 que se da 2 veces; 4 que se da 3 veces; 5 que se da 4 veces; 6 que se da 3 veces; 7 que se da 2 y 8 que se da 1 vez. De otra parte, el valor 2 se da cuando $X = 1$ y $Y = 1$ (y sólo en este caso), así que

$P[(X + Y) = 2] = P[X = 1, Y = 1] = \frac{1}{16}$, de acuerdo con el espacio muestral dado antes.

Igualmente, $X + Y = 3$ se da cuando ocurre una de las dos situaciones $[X = 1, Y = 2]$ o $[X = 2, Y = 1]$ y así $P[(X + Y) = 3] = \frac{2}{16}$

Resumiendo tenemos que la distribución de $X + Y$ es como se indica en la tabla 7.3.

Tabla 7.3 *Distribución de X + Y.*

$X + Y$	2	3	4	5	6	7	8
$P[(X + Y) = z]$:	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

Vamos ahora a calcular $E(X + Y)$. Para ello nos remitimos a la fórmula (5-4) y tenemos,

$$E(X + Y) = 2\left(\frac{1}{16}\right) + 3\left(\frac{2}{16}\right) + 4\left(\frac{3}{16}\right) + 5\left(\frac{4}{16}\right) + 6\left(\frac{3}{16}\right) + 7\left(\frac{2}{16}\right) + 8\left(\frac{1}{16}\right) =$$

$$= \frac{2 + 6 + 12 + 20 + 18 + 14 + 8}{16} = \frac{80}{16} = 5$$

Si sólo tomamos en cuenta los valores de la variable X, su distribución es como se muestra en la tabla 7.4.

Tabla 7.4 *Distribución de X.*

X	1	2	3	4
$P[X = x]$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

y su valor esperado es $E(X) = 1\left(\frac{1}{4}\right) + 2\left(\frac{1}{4}\right) + 3\left(\frac{1}{4}\right) + 4\left(\frac{1}{4}\right) = \frac{1 + 2 + 3 + 4}{4} = \frac{10}{4} = \frac{5}{2}$

Una tabla similar para Y nos mostrará que $E(Y)$ es también $\frac{5}{2}$

De lo anterior se tiene que $E(X) + E(Y) = \frac{5}{2} + \frac{5}{2} = 5$ que coincide con $E(X + Y)$.

Esto es, $E(X + Y) = E(X) + E(Y)$

En general se tiene que, si $X_1, X_2, \ldots, X_n$ son n variables aleatorias entonces,

$$E(X_1 + X_2 + \ldots + X_n) = E\left(\sum_{i=1}^{n} X_i\right) = E(X_1) + E(X_2) + \ldots + E(X_n) = \sum_{i=1}^{n} E(X_i) \qquad (7\text{-}5)$$

La expresión (7-5) es un caso particular de (5-6).

Tomemos ahora en consideración la variable producto. Al hacer un análisis similar al de la suma, se obtiene la tabla 7.5 que representa la distribución de XY.

Tabla 7.5 *Distribución de XY.*

XY	1	2	3	4	6	8	9	12	16
$P[(XY) = z]$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

$$E[(XY)] = 1\left(\tfrac{1}{16}\right) + 2\left(\tfrac{2}{16}\right) + 3\left(\tfrac{2}{16}\right) + 4\left(\tfrac{3}{16}\right) + 6\left(\tfrac{2}{16}\right) + 8\left(\tfrac{2}{16}\right) + 9\left(\tfrac{1}{16}\right) + 12\left(\tfrac{2}{16}\right) + 16\left(\tfrac{1}{16}\right)$$

$$= \frac{1 + 4 + 6 + 12 + 12 + 16 + 9 + 24 + 16}{16} = \frac{82}{16} = 6.25$$

Observe que $E(XY) = E(X)E(Y)$. Esto no siempre es así, como se verá en uno de los problemas propuestos en los ejercicios.

¿Qué decir de la independencia de X y Y?
En respuesta a este interrogante debemos determinar si se cumple $P[X = x, Y = y] = P[X = x]\,P[Y = y]$

Para tal propósito, tengamos en cuenta que los valores $P[X = x]$ están dados en la tabla 7.4 y que los valores de Y se obtienen de manera similar; sólo nos falta obtener la distribución (conjunta) de X y Y. Esta distribución está descrita en el cuadro 3, en donde los números de la primera columna corresponden a los resultados del primer tetraedro y los de la primera fila al segundo.
Los valores del interior del cuadro se obtienen con base en el espacio muestral,
$S = \{1\text{-}1, 1\text{-}2, 1\text{-}3, 1\text{-}4, \ldots, 4\text{-}1, 4\text{-}2, 4\text{-}3, 4\text{-}4\}$

X \ Y	1	2	3	4
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
2	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
3	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
4	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

Cuadro 3 *Distribución conjunta de X, Y.*

Los valores del cuadro se interpretan así: $P[X = 1, Y = 1] = \frac{1}{16}$, $P[X = 1, Y = 2] = \frac{1}{16}$ y así sucesivamente.

Por otra parte, $P[X = 1] = \frac{1}{4}$, $P[Y = 1] = \frac{1}{4}$ y de esta forma

$P[X = 1, Y = 1] = \frac{1}{16} = (\frac{1}{4})(\frac{1}{4}) = P[X = 1]\ P[Y = 1]$. Igual relación se cumple para las restantes parejas de números. Por tanto, X y Y son independientes. Cuando X y Y son independientes se puede demostrar que $E(XY) = E(X)E(Y)$[1] (7-6) que es lo que ocurre en el ejercicio que hemos venido desarrollando.

Ya hemos visto que $E(X + Y) = E(X) + E(Y)$, pero ¿se cumple también que $V(X + Y) = V(X) + V(Y)$? Veamos.

En el capítulo 5, se estableció que si se tiene una variable aleatoria X entonces la varianza de X está dada por $V(X) = E(X^2) - (E[X])^2$. Aplicada esta fórmula para el caso de la suma $X + Y$ se tiene $V(X + Y) = E[(X + Y)^2] - (E[X + Y])^2$

Al desarrollar por separado se tiene:

$E[(X + Y)^2] = E(X^2 + 2XY + Y^2) = E(X^2) + 2E(XY) + E(Y^2)$ (aplicando 5-6)

$(E[X + Y])^2 = (E[X] + E[Y])^2 = (E[X])^2 + 2E[X]\ E[Y] + (E[Y])^2$

De lo anterior se tiene:

$V(X + Y) = E(X^2) + 2E(XY) + E(Y^2) - (E[X])^2 - 2E[X]\ E[Y] - (E[Y])^2 =$

$(E(X^2) - (E[X])^2 + (E(Y^2) - (E[Y])^2 + 2[E(XY) - E(X)E(Y)]$ (7-7)

Por tanto, $V(X + Y) = V(X) + V(Y) + 2\ [E(XY) - (E(X)E(Y)]$.

Así que podemos afirmar que $V(X + Y) = V(X) + V(Y)$ cuando el último término del lado derecho de (7-7) sea cero. Pero esto ocurre cuando $E(XY) - E(X)\ E(Y) = 0$, esto es, si $E(XY) = E(X)\ E(Y)$, lo cual se da cuando X y Y son independientes. En resumen,

Si X y Y son variables aleatorias independientes entonces $V(X + Y) = V(X) + V(Y)$.[2]

Este resultado se puede generalizar y se tiene:

Si $X_1, X_2, \ldots, X_n$ son n variables aleatorias independientes, entonces

$$V(X_1 + X_2 + \ldots + X_n) = V(\sum_{i=1}^{n} X_i) = V(X_1) + V(X_2) + \ldots + V(X_n) = \sum_{i=1}^{n} V(X_i) \qquad (7\text{-}8)$$

La expresión dada por (7-7) corresponde a la varianza de una suma en general y la diferencia $E(XY) - E(X)E(Y)$ se llama covarianza de X y Y y se denota $\text{Cov}(X, Y)$. Esto es,

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) \qquad (7\text{-}9)$$

La covarianza es de gran importancia en el estudio de relación entre variables.

Estos conceptos y propiedades que hemos deducido para variables aleatorias discretas también son válidos para el caso de variables aleatorias continuas.

7.3 MUESTRA ALEATORIA. ESTADÍSTICAS

Como ya hemos comentado antes cuando tomamos en cuenta cierta característica para su estudio, esta característica determina o define la variable aleatoria. Ésta se

[1] En el suplemento III se encuentra una demostración de esta igualdad.

[2] En el suplemento III se encuentra una demostración un poco diferente.

comportará de acuerdo con cierto patrón probabilístico (llámese binomial, Poisson, normal, etc.) y de esta manera hablamos de la distribución de la variable o de la población.

Supongamos ahora que hemos convenido en denotar X la variable (la característica) que pretendemos estudiar y que vamos a hacer n observaciones en la población respectiva. Estas observaciones serán datos concretos una vez que hayamos llevado a cabo el acto físico de tomarlas; antes, sólo podemos considerar valores posibles de acuerdo con la distribución de X, esto es, la respuesta es aleatoria. Por ello cada una de estas observaciones que luego se materializarán, las denotaremos $X_1, X_2, \ldots, X_n$ y se consideran n "representaciones" de la variable X, y por tanto con la misma distribución de X. Si además las variables $X_1, X_2, \ldots, X_n$ se consideran independientes, tenemos lo que se llama una *muestra aleatoria*.

Una muestra aleatoria de una población X *es una sucesión* $X_1, X_2, \ldots, X_n$ *de* n *variables aleatorias independientes e igualmente distribuidas (tienen la misma función de densidad).*

Nota. Cuando se dice que las variables $X_1, X_2, \ldots, X_n$ tienen la misma distribución debe entenderse que tienen la misma función de densidad y por tanto la misma media y la misma varianza. Con esta sucesión de variables aleatorias $X_1, X_2, \ldots, X_n$ se pueden llevar a cabo operaciones aritméticas tales como:

$$Y = \frac{\sum_{i=1}^{n} X_i}{n}, \qquad U = \frac{\sum_{i=1}^{n} X^2_i}{n}, \qquad V = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n}, \qquad W = \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{n}$$

Algunas de estas fórmulas pueden ser de tal índole que sólo sean desconocidos los valores $X_1, X_2, \ldots, X_n$, esto es, no contienen constantes desconocidas. Cuando esto ocurre, a tales fórmulas se les llama *estadísticas*.

Una estadística es cualquier fórmula matemática que relaciona las variables de una muestra aleatoria $X_1, X_2, \ldots, X_n$ *y que no incluye constantes desconocidas.*
Así por ejemplo,

$$Y = \frac{\sum_{i=1}^{n} X_i}{n}, \qquad U = \frac{\sum_{i=1}^{n} X^2_i}{n}, \qquad V = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n}, \text{ son estadísticas.}$$

Pero, en cambio, si μ es desconocida entonces, $W = \dfrac{\sum_{i=1}^{n} (X_i - \mu)^2}{n}$, no es una estadística.

El proceso inferencial se lleva a cabo utilizando las estadísticas (variables) como medio para tal fin y son las de mayor uso las denominadas media y varianza muestral que se denotan y definen así:

Media muestral: $\overline{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ *Varianza muestral:* $S^2 = \dfrac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{(n-1)}$

Las estadísticas son de por sí variables aleatorias; por ello es de esperarse que tengan asociadas distribuciones y tenemos:

La distribución muestral o de muestreo de una estadística T es la distribución de probabilidad de T, tomada ésta como una variable aleatoria.
Así por ejemplo, cuando decimos que $\overline{X}$ (que es una estadística) tiene distribución normal nos estamos refiriendo a una distribución muestral.

Es claro que cuando tomamos en consideración una variable X, por lo general, no va a ser posible observar todos los valores de la variable y sólo conoceremos algunos de sus valores (los proporcionados por los datos muestrales). Sin embargo, al menos en sentido teórico, podemos pensar en ciertos números resultantes de ciertas operaciones en donde intervienen todos los valores de la variable con sus respectivas probabilidades (la media y la varianza de la variable son números de este tipo). Como tales números resultan por consideración de todos los valores de la variable conjuntamente con sus probabilidades respectivas, permanecerán inalterados y en cierta forma caracterizarán la distribución de la variable; llegamos de esta manera al concepto de *parámetro*.

Un parámetro es una caracterización numérica de la distribución de la población de forma que describe total o parcialmente la función de densidad de la variable aleatoria de interés.
La media y la varianza de una variable aleatoria con distribución normal son parámetros.

Debe quedar clara la diferencia y relación que existe entre una estadística y un parámetro. La estadística se calcula de acuerdo con las variables aleatorias de la muestra, por consiguiente cambia de muestra a muestra, pero sigue cierta ley de probabilidad, lo que constituye la distribución muestral. El parámetro es una característica de la población y como tal permanece constante (al menos dentro de cierto límite de tiempo) y generalmente es desconocido. Pero a cada parámetro casi siempre se le puede asociar una estadística mediante la cual podemos obtener alguna información acerca del parámetro desconocido. Cómo se realiza este proceso es la esencia de la inferencia estadística en cualquiera de sus ramas (estimación y prueba de hipótesis).

Ejercicios 7.1

1. Explique la diferencia que existe entre parámetro y estadística.

2. Defina distribución muestral.

3. Si X y Y son variables aleatorias, ¿en qué caso se cumple $V(X + Y) = V(X) + V(Y)$?

4. Si X y Y son variables aleatorias independientes, demuestre que $V(X - Y) = V(X) + V(Y)$.

5. Suponga que la variable X asume los valores 1, 2, 3 con la misma probabilidad; la variable Y asume los valores 0, 1 con la misma probabilidad. Si X y Y son independientes y $W = 2X + 3Y$, compruebe que $\sigma^2_w = 4\sigma^2_x + 9\sigma^2_y$.

6. Dadas las variables aleatorias X y Y del problema 5 y si $Z = 2X - Y$, verifique que

$$\sigma^2_z = 4\sigma^2_x + \sigma^2_y.$$

7. Se lanzan un par de tetraedros. Sean las variables, X = número de puntos que muestra la cara sobre la cual se apoya el primer tetraedro, Y = suma de los puntos que muestran las dos caras sobre las cuales se apoyan los dos tetraedros.

a) Halle la distribución de X
b) Halle la distribución de Y.
c) Halle la distribución de XY
d) Halle la distribución conjunta de X y Y
e) Compruebe que $E(XY) = E(X)E(Y)$

8. De una población normal de media $\mu_x = 5$, se tomó una muestra de tamaño 10 y se obtuvo una media muestral $\bar{x} = 4.5$. Identifique en este problema el valor del parámetro y el valor de la estadística.

9. ¿Cuál es la relación que existe entre parámetro y estadística?

10. Proponga una variable aleatoria X como tema de estudio; ¿qué parámetro asociado tendría esta variable?

11. Defina la distribución muestral.

12. Sean X y Y dos variables aleatorias, ¿en qué caso se cumple $E(X + Y) = E(X) + E(Y)$?

7.4 DISTRIBUCIÓN DE LA MEDIA MUESTRAL

Como ya se señaló, la media muestral $\bar{X}$ al ser una estadística, es por naturaleza una variable aleatoria y como tal tiene una distribución y también tendrá una media (valor esperado) y una varianza.

A partir de la tabla 7.2 y teniendo en cuenta la definición de valor esperado discutida en el capítulo 5, tenemos que el valor esperado para $\bar{X}$ está dado por

$$E(\bar{X}) = \mu_{\bar{x}} = 2\left(\frac{1}{16}\right) + 3\left(\frac{2}{16}\right) + 4\left(\frac{3}{16}\right) + 5\left(\frac{4}{16}\right) + 6\left(\frac{3}{16}\right) + 7\left(\frac{2}{16}\right) + 8\left(\frac{1}{16}\right) =$$

$$= \frac{2 + 6 + 12 + 20 + 18 + 14 + 8}{16} = \frac{80}{16} = 5$$

Se sabe que la media de la población a partir de la cual se toman las muestras está dada por $\mu_x = \dfrac{2 + 4 + 6 + 8}{4} = \dfrac{20}{4} = 5$. Así que la media de $\bar{X}$ es igual a la media de X. Esto es, $\mu_{\bar{X}} = \mu_X$. Este resultado no es una particularidad del presente problema, sino que es una propiedad entre estas dos medias. En efecto, supongamos que $X_1, X_2, \ldots, X_n$ es una muestra aleatoria y $\bar{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ es la media muestral

entonces $E(\bar{X}) = E\left(\dfrac{\sum_{i=1}^{n} X_i}{n}\right) = \dfrac{1}{n} E\left(\sum_{i=1}^{n} X_i\right) = \dfrac{1}{n} E(X_1 + X_2 + \ldots + X_n) =$

$\dfrac{1}{n}[E(X_1) + E(X_2) + \ldots + E(X_n)] = \dfrac{1}{n}[E(X) + E(X) + \ldots + E(X)] =$

$\dfrac{1}{n}[nE(X)] = E(X)$

En cuanto a la varianza de $\overline{X}$, para los datos de la tabla 7.2 se tiene,

$$V(\overline{X}) = E(\overline{X}^2) - [E(\overline{X})]^2$$

$$E(\overline{X}^2) = 2^2\left(\tfrac{1}{16}\right) + 3^2\left(\tfrac{2}{16}\right) + 4^2\left(\tfrac{3}{16}\right) + 5^2\left(\tfrac{4}{16}\right) + 6^2\left(\tfrac{3}{16}\right) + 7^2\left(\tfrac{2}{16}\right) + 8^2\left(\tfrac{1}{16}\right) =$$

$$= \frac{4 + 18 + 48 + 100 + 108 + 98 + 64}{16} = \frac{440}{16} = 27.5, \text{ así que}$$

$$V(\overline{X}) = 27.5 - 5^2 = 27.5 - 25 = 2.5$$

Por otra parte, $V(X) = E(X^2) - [E(X)]^2$

$$E(X^2) = 2^2\left(\tfrac{1}{4}\right) + 4^2\left(\tfrac{1}{4}\right) + 6^2\left(\tfrac{1}{4}\right) + 8^2\left(\tfrac{1}{4}\right) = \frac{4 + 16 + 36 + 64}{4} = \frac{120}{4} = 30$$

$$V(X) = 30 - 5^2 = 30 - 25 = 5$$

Al comparar los dos resultados observamos que $V(\overline{X}) = 2.5 = \frac{5}{2} = \frac{V(X)}{n}$.

Esta relación entre $V(\overline{X})$ y $V(X)$ siempre se cumple cuando $\overline{X}$ se trata de la media muestral, como se demuestra a continuación.

$$V(\overline{X}) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \left(\frac{1}{n}\right)^2 V\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} V(X) = \frac{1}{n^2}(nV(X)) = \frac{V(X)}{n} \quad \text{por (7-8)}$$

Resumiendo las dos propiedades hasta aquí discutidas para la media muestral $\overline{X}$, tenemos el siguiente teorema:

Teorema 7.1 Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria proveniente de una población de media μ_x y varianza σ^2_x. Si $\overline{X}$ es la media muestral,

entonces $E(\overline{X}) = \mu_{\overline{x}} = E(X) = \mu_x$ y $V(\overline{X}) = \sigma^2_{\overline{X}} = \dfrac{\sigma^2_X}{n}$ (7-10)

A partir de la varianza de $\overline{X}$, $\sigma^2_{\overline{X}} = \dfrac{\sigma^2_X}{n}$ se obtiene la desviación estándar de $\overline{X}$, dada por $\sigma_{\overline{x}} = \dfrac{\sigma_X}{\sqrt{n}}$ y que recibe el nombre de *error estándar de la media.*

El error estándar, como nos lo muestra la fórmula, es directamente proporcional a σ_x e inversamente proporcional a la raíz cuadrada del tamaño de la muestra, por consiguiente, dada σ_x, se puede controlar el valor de $\sigma_{\overline{x}}$ al controlar el tamaño de la muestra, sea aumentándola o disminuyéndola.

Ahora bien, si el tamaño de la muestra es 1, la varianza de la media es exactamente la de la población. Si n es igual a dos, entonces $\sigma^2_{\overline{x}}$ es la mitad de σ^2_x. Si n es igual a 100, $\sigma^2_{\overline{x}}$ será la centésima parte de σ^2_x. Si n se hace infinitamente grande, $\sigma^2_{\overline{x}}$ tiende a cero, lo cual significa que la medición muestral es idéntica a la media de la población. En general, cuanto mayor sea el valor n, más probable es que $\overline{X}$ se aproxime indefinidamente a la media μ de la población. Al tomar n suficientemente grande, se puede hacer que la probabilidad de que $\overline{X}$ difiera de μ en un valor d, sea tan grande como se desee. Esta relación es conocida como ley de los *grandes números.*

Esta ley puede explicarse mediante la desigualdad de Shebyshev (que es la versión teórica de la regla del mismo nombre).

La desigualdad de Shebyshev dice:

"Si X es una variable aleatoria de media μ y varianza $\sigma^2_x \Longrightarrow P[|X - \mu_x| \leq k\sigma_x] \geq 1 - \frac{1}{k^2}$."[3]

La cual se interpreta al decir "la probabilidad de que X quede dentro de k desviaciones estándares a partir de la media es por lo menos $1 - \frac{1}{k^2}$, que aplicado a la media muestral $\overline{X}$ nos queda:

$$P[|\overline{X} - \mu| \leq k\left(\frac{\sigma_x}{\sqrt{n}}\right)] \geq 1 - \frac{1}{k^2}$$

al ser $k\left(\frac{\sigma_X}{\sqrt{n}}\right) = d$ del que se habló antes. Sea d un número positivo pequeño cualquiera.

Si la muestra es suficientemente grande, se puede hacer la probabilidad de que $\overline{X}$ esté menos de d de μ, tan próximo a 1 como se desee.
Por ejemplo, dado $\sigma_x = 10$, si se quiere que la media muestral no esté a más de 5 unidades de la media de la población con probabilidad del 99% por lo menos, habría que hacer el tamaño de la muestra igual a 400, lo cual se calcula como sigue. Como

$$1 - \frac{1}{k^2} = 0.99 \Longrightarrow k = 10$$

Por otra parte $d = 5 = 10\left(\frac{10}{\sqrt{n}}\right) \Longrightarrow 5\sqrt{n} = 100 \Longrightarrow \sqrt{n} = 20 \Longrightarrow n = 400$

Hasta aquí hemos estudiado las propiedades de la media muestral $\overline{X}$ sin hacer consideración alguna de la distribución de la variable en cuestión. Ahora nos proponemos estudiar el comportamiento de $\overline{X}$ cuando X se distribuye normalmente.

7.4.1 Distribución de $\overline{X}$ en una población normal con media y varianza conocidas

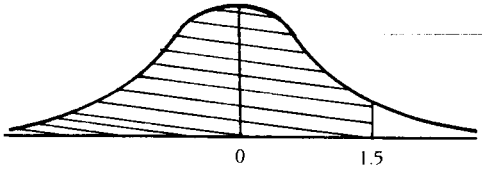
Un resultado que se demuestra en estudios de estadística matemática es el siguiente:

Teorema 7.2 Si $X_1, X_2, \ldots, X_n$ es una muestra aleatoria proveniente de una población normal de media μ y varianza $\sigma^2_x \Longrightarrow$ la variable $Y = \sum_{i=1}^{n} a_i X_i$ tiene distribución normal con media $\mu_y = (\sum_{i=1}^{n} a_i)\mu$ y varianza $\sigma^2_y = (\sum_{i=1}^{n} a_i^2)\sigma^2$

Así por ejemplo, si se tiene una muestra aleatoria X_1, X_2, X_3 proveniente de una población normal de media 10 y varianza 9, la variable $Y = X_1 + 2X_2 - X_3$ tiene distribución normal con media $\mu_y = 2(10) = 20$ y varianza $\sigma^2_y = (6)(9) = 54$.
Del anterior teorema se tiene que si requerimos el cálculo de $P[Y < 31]$, por ejemplo. Entonces, partiendo del hecho $Y \sim N(20, 54)$ y mediante la estandarización de Y tenemos con utilización de la tabla II.

$$P[\frac{(Y - 20)}{\sqrt{54}} < \frac{(31 - 20)}{\sqrt{54}}] = P[Z < \frac{11}{\sqrt{54}}] = P[Z < 1.5] = 0.9332$$

[3] En el suplemento IV se da una demostración de esta desigualdad.



Del teorema 7.2, se deduce el teorema que sigue:

Teorema 7.3 Si $X_1, X_2, \ldots, X_n$ es una muestra aleatoria proveniente de una población con distribución normal de media μ y varianza $\sigma^2 \Longrightarrow \overline{X} = (\sum_{i=1}^{n} X_i)/n$ tiene distribución normal con media μ y varianza $\frac{\sigma^2}{\sqrt{n}}$. Esto es, $\overline{X} \sim N(\mu, \frac{\sigma^2}{\sqrt{n}})$

Es claro que si $\overline{X} \sim N(\mu, \sigma^2/\sqrt{n}) \Longrightarrow$ La variable $Z = \frac{(\overline{X} - \mu)}{\sigma/\sqrt{n}} =$

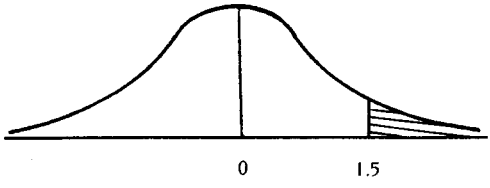
$$= \frac{\sqrt{n}(\overline{X} - \mu)}{\sigma} \sim N(0, 1).$$

Suponga que una máquina dispensadora de bebidas gaseosas la cantidad que envasa es una variable aleatoria X que tiene distribución normal con media $\mu = 10$ onzas y desviación estándar de $\sigma = 1$. Y nos proponemos hacer 25 mediciones del líquido dispensado.

a) Exprese el significado de $\overline{X}$
b) ¿Qué distribución tiene $\overline{X}$?
c) ¿Cuál es la probabilidad de que $\overline{X}$ sea por lo menos 10.3?
En primer lugar tenemos a) $\overline{X}$ = cantidad promedio de gaseosa dispensada en las 25 mediciones.
b) De acuerdo con el teorema 7.3, $\overline{X}$ tiene distribución normal con media $\mu_{\overline{x}} = \mu_x = 10$ y varianza $\sigma^2_{\overline{x}} = \frac{\sigma^2}{n} = \frac{1}{25}$

c) $P[\overline{X} \geq 10.3] = P[\frac{\overline{X} - \mu}{1/5} \geq \frac{10.3 - 10}{1/5}] = P[Z \geq 1.5] = 1 - P[Z \leq 1.5]$

$= 1 - 0.9332 = 0.0668$



Otro problema. El tiempo que un cajero de banco tarda en atender a los clientes que se acercan a su ventanilla para solicitar servicio es una variable aleatoria distribuida normalmente. Se asegura que el tiempo promedio que tarda este cajero en atender

el público es de 3.5 minutos con una desviación estándar de 1.2 minutos. Un observador inquieto tomó el tiempo a 36 clientes y obtúvo un promedio $\bar{x} = 4.3$. ¿Qué puede decirse del promedio $\mu = 3.5$ de acuerdo con este valor muestral?

En este problema podemos ver algo del proceso inferencial. En la medida en que debemos tomar una decisión acerca de un valor de parámetro, la media en este caso, basado en un dato muestral.

El procedimiento a desarrollar consiste de los siguientes pasos:

1°. Admitimos que lo afirmado acerca del tiempo requerido por el cajero es cierto. Esto se plantea de la siguiente forma: Consideremos la variable, X = tiempo que tarda el cajero en atender cada cliente. Ésta, de acuerdo con las condiciones del problema, tiene distribución normal con media $\mu = 3.5$ y varianza $\sigma^2 = (1.20)^2$ que gráficamente se ilustra como sigue:

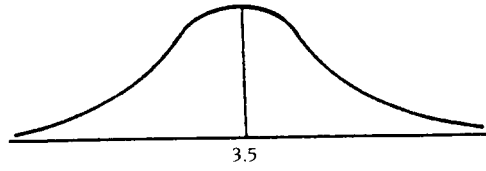


Figura 7.2 *Distribución de X.*

2°. Localizamos sobre el eje horizontal el valor muestral, 4.3 en este caso

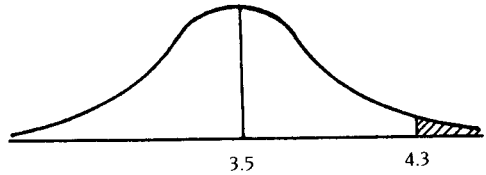


Figura 7.3 *Distribución de* $\bar{X}$.

3°. Calculamos la probabilidad (área) de que $\bar{X}$ sea mayor o igual a 4.3

$$P[\bar{X} \geq 4.3] = P[\frac{\bar{X} - 3.5}{1.2/\sqrt{36}} \geq \frac{4.3 - 3.5}{0.2}] = P[Z \geq 0.8/0.2]$$
$$= P[Z \geq 4] = 0$$

¿Cómo interpretar este valor de probabilidad?

Esta probabilidad de cero se interpreta diciendo que el valor muestral está por fuera de la curva que representa $\bar{X}$. Como esta curva es el resultado de la consideración que hemos hecho sobre X, entonces estamos frente a una inconsistencia entre lo supuesto y la realidad (el dato muestral).

Como el dato muestral es incuestionable, el único camino que nos queda para romper la inconsistencia es admitir que lo supuesto no es correcto. Tal incorrección puede darse por muchas razones. Pero, como lo que se sometió a discusión fue la media, suponemos que sobre lo demás no hay duda y así tenemos como única salida admitir que el valor de la media de 3.5 minutos no es correcto. Para el mismo problema suponga que el valor muestral fue de $\bar{x} = 3.8$. En este caso, el paso 2° queda representado en la figura 7.4 y calculamos el área (probabilidad) a la derecha de 3.8.

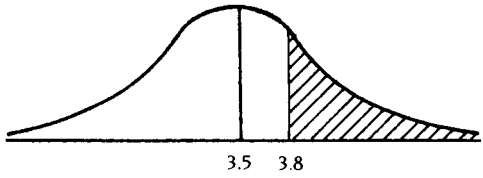


Figura 7.4 *Distribución de* $\overline{X}$.

$$P[\overline{X} \geq 3.8] = P[\frac{\overline{X} - 3.5}{1.2/\sqrt{36}} \geq \frac{3.8 - 3.5}{0.2}] = P[Z \geq 1.5] = 0.068$$

Lo que indica una porción de área del 6.68%.

En este caso, sí admitimos que el valor muestral forma parte de la región limitada por la curva y decimos que no hay evidencia para rechazar el valor $\mu = 3.5$.

Nota. Para que un valor $\overline{x}$ se considere dentro de la región debe dejar un área de cola con un mínimo del 5%.

Ahora supongamos que el valor muestral fue de 3.

En estas circunstancias el paso 2° toma la forma que se presenta en la figura 7.5.

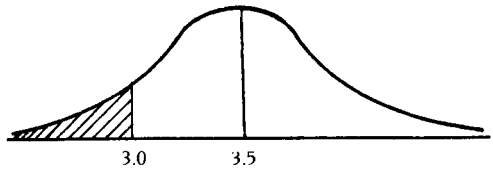


Figura 7.5 *Distribución de* $\overline{X}$.

El área a analizar ahora es la que queda a la izquierda de 3.5 y así calculamos $P[\overline{X} \leq 3]$

$$P[\frac{\overline{X} - 3.5}{0.2} \leq \frac{3 - 3.5}{0.2}] = P[Z \leq -2.5] = 1 - P[Z \leq 2.5]$$

$$= 1 - 0.9938 = 0.0062$$

y así el valor $\mu = 3.5$ parece no ser correcto.

Finalmente, algunos comentarios acerca del comportamiento de $\overline{X}$ que en poblaciones normales pueden resultar oportunos. Al tener $\overline{X}$ también distribución normal, la forma general de la curva de X y $\overline{X}$ es la misma. Pero la de $\overline{X}$ es más concentrada alrededor de la media de la población. Si $n = 1$, los valores de $\overline{X}$ serán los mismos de X. Si el tamaño de la muestra es mayor, el error estándar de $\overline{X}$, $\sigma/\sqrt{n}$, será menor y, en consecuencia la cresta de la función de densidad será más alta.

7.4.2 Teorema del límite central

En la sección precedente hemos estudiado el comportamiento de $\overline{X}$ en poblaciones normales y se obtuvo como conclusión que $\overline{X}$ tiene también distribución normal. Más concretamente, la variable $Z = \frac{\sqrt{n}(\overline{X} - \mu)}{\sigma}$ tiene distribución normal estándar. Sin embargo, este resultado es aproximadamente correcto aun en poblaciones no normales, como se establece en el siguiente teorema, llamado *teorema del límite central.*

Teorema 7.4 Teorema del límite central.
Si $\overline{X}$ es la media de una muestra aleatoria de tamaño n que se toma de una población con media μ y varianza finita σ^2, entonces la variable $Z_n = \dfrac{\sqrt{n}(\overline{X} - \mu)}{\sigma}$ tiende a la normal estándar a medida que n tiende a infinito.

La importancia de este resultado radica en que nos proporciona un medio para trabajar con $\overline{X}$ aun si desconocemos qué distribución tiene X (la variable objeto de estudio), de ahí su gran utilidad práctica.

El teorema del límite central puede describirse de otra forma que puede resultar más clara para su interpretación.

"Si X es una variable aleatoria de media μ y varianza σ^2, la distribución muestral de la media $\overline{X}$ de una muestra aleatoria de tamaño n es aproximadamente normal con media μ y varianza σ^2/n si n es suficientemente grande". Simbólicamente se escribe $\overline{X} \sim N(\mu, \sigma^2/n)$.

Cabe preguntarse, ¿a partir de qué valor n se puede considerar "suficientemente" grande? En muchos casos, la tendencia de $\overline{X}$ a distribuirse en forma normal es bastante acentuada, aun cuando las muestras sean de tamaño moderado. Si bien se prefiere $n \geq 100$, en la mayoría de las aplicaciones se considera suficiente una muestra de tamaño 30 o más para permitir el empleo de la distribución normal a fin de encontrar las probabilidades asociadas a $\overline{X}$.

La demostración del teorema del límite central escapa al alcance de este libro; pero con el sencillo ejemplo que sigue, se mostrará la tendencia de la distribución de $\overline{X}$ de acercarse cada vez más a la distribución normal a medida que crece el tamaño de la muestra. El ejercicio que desarrollaremos se hará al tomar los datos del problema que se consideró al comienzo de este capítulo.

Se trata de una variable aleatoria X que toma los valores 2, 4, 6, 8. La distribución de esta variable es como se indica en la tabla que sigue y su representación gráfica es como se señala en la figura 7.6. Es la distribución de $\overline{X}$ cuando el tamaño de la muestra es igual a 1.

Distribución de $\overline{X}$

x	$P(x)$
2	$\frac{1}{4}$
4	$\frac{1}{4}$
6	$\frac{1}{4}$
8	$\frac{1}{4}$

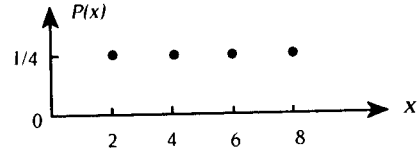
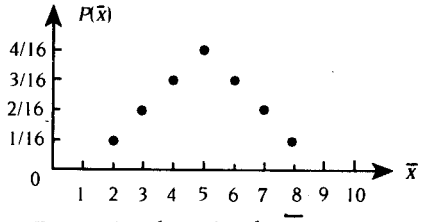


Figura 7.6 *Distribución de* $\overline{X}$.

Si se toman de esta población muestras con repetición de tamaño 2, o sea X_1, X_2, la suma de los valores de las dos observaciones para toda muestra posible aparece en las tablas que se dan a continuación acompañadas de la distribución muestral de $\overline{X}$. La figura de abajo nos muestra la representación gráfica de $\overline{X}$; en ella se puede notar cómo la distribución se acerca más a la normal.

+	2	4	6	8
2	4	6	8	10
4	6	8	10	12
6	8	10	12	14
8	10	12	14	16

$X_1 + X_2$



Media muestral $\bar{x}$	*Probabilidad* $P(\bar{x})$
2	$\frac{1}{16}$
3	$\frac{2}{16}$
4	$\frac{3}{16}$
5	$\frac{4}{16}$
6	$\frac{3}{16}$
7	$\frac{2}{16}$
8	$\frac{1}{16}$
Total	1.0

Distribución de $\bar{X}$.

Figura 7.7 *Distribución de* $\bar{X}$.

Si se aumentara el tamaño de la muestra habría un mayor número de muestras posibles y de medias muestrales y los puntos del gráfico estarían más cercanos a sus puntos adyacentes de los que están en la figura 7.7. Por ejemplo, si $n = 3$, habrá 10 diferentes medias muestrales posibles, como se muestra en la tabla 7.6.

La distribución y gráfica de $\bar{X}$ para $n = 3$ es como se indica en la tabla siguiente y en la figura 7.8.

Media muestral $\bar{x}$	*Probabilidad* $P(\bar{x})$
2	$\frac{1}{64}$
$2\frac{2}{3}$	$\frac{3}{64}$
$3\frac{1}{3}$	$\frac{6}{64}$
4	$\frac{10}{64}$
$4\frac{2}{3}$	$\frac{12}{64}$
$5\frac{1}{3}$	$\frac{12}{64}$
6	$\frac{10}{64}$
$6\frac{2}{3}$	$\frac{6}{64}$
7	$\frac{3}{64}$
8	$\frac{1}{64}$
Total	1.0

Distribución de $\bar{X}$.

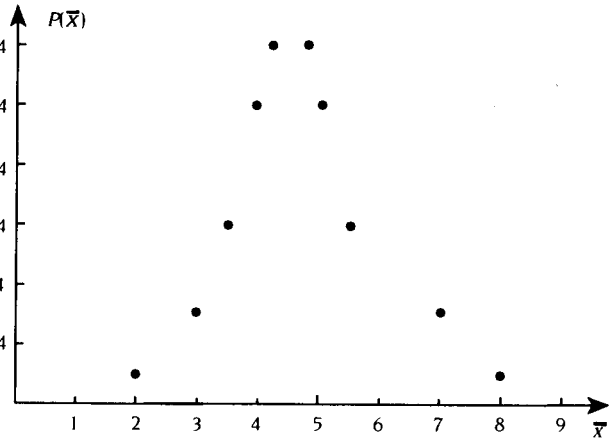


Figura 7.8 *Distribución de* $\bar{X}$.

Consideremos otro ejemplo. Para cierta prueba de aptitud se sabe con base en la experiencia que el número de aciertos es en promedio 500 con una desviación estándar de 60. Si se aplica esta prueba a 100 personas seleccionadas al azar, halle

Tabla 7.6 *Posibles muestras y medias con* $n = 3$, *cuando* X *tiene el conjunto de valores posibles* {2, 4, 6, 8}.

Muestras posibles X_1	X_2	X_3	Suma $\sum_{i=1}^{3} X_i$	Media $\bar{X}$	Número de muestras
2	2	2	6	2	1
2	2	4			
2	4	2	8	$2\frac{2}{3}$	3
4	2	2			
2	4	4			
4	2	4			
4	4	2	10	$3\frac{1}{3}$	6
2	2	6			
2	6	2			
6	2	2			
4	4	4			
2	4	6			
2	6	4			
4	2	6			
4	6	2	12	4	10
6	2	4			
6	4	2			
2	2	8			
2	8	2			
8	2	2			
2	6	6			
6	2	6			
6	6	2			
2	4	8			
2	8	4			
4	2	8	14	$4\frac{2}{3}$	12
4	8	2			
8	2	4			
8	4	2			
4	4	6			
4	6	4			
6	4	4			
2	6	8			
2	8	6			
6	2	8			
6	8	2			
8	2	6			
8	6	2	16	$5\frac{1}{3}$	12
4	6	6			
6	4	6			
6	6	4			
4	4	8			
4	8	4			
8	4	4			
2	8	8			
8	2	8			
8	8	2			
4	6	8			
4	8	6	18	6	10
6	4	8			
6	8	4			
8	4	6			
8	6	4			
6	6	6			
4	8	8			
8	4	8			
8	4	4	20	$6\frac{2}{3}$	6
6	6	8			
6	8	6			
8	6	6			
6	8	8			
8	6	8	22	$7\frac{1}{3}$	3
8	8	6			
8	8	8	24	8	1

la probabilidad de que tengan un promedio de aciertos, a) mayor de 512; b) menor de 496.

En primer lugar precisamos la variable que vamos a analizar.

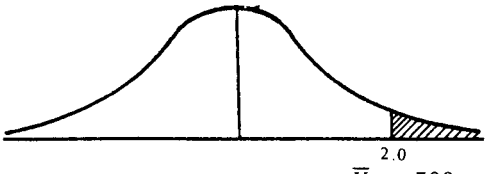
X = Número de aciertos que tiene la persona.

La distribución de esta variable no es conocida. Pero, como se trata de responder una probabilidad asociada con $\bar{X}$, podemos fundamentarnos en el teorema del límite central para dar respuestas a los interrogantes planteados.

De acuerdo con el teorema del límite central, $\overline{X} \sim N(500, (60)^2/100)$ y así para la parte a).

$$P[\overline{X} \geq 512] = P\left[\frac{\overline{X} - 500}{6} \geq \frac{512 - 500}{6}\right] = P[Z \geq 2] = 1 - P[Z \leq 2] =$$

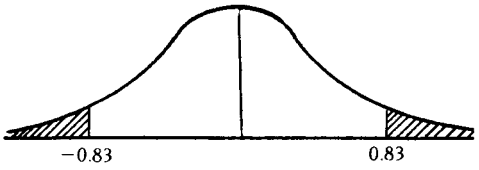
$$= 1 - 0.9772 = 0.0228$$



Para el literal b) se tiene, $P[\overline{X} < 495] = P\left[\frac{\overline{X} - 500}{6} < \frac{495 - 500}{6}\right] =$

$$= P\left[Z < \frac{-5}{6}\right] = P[Z < -0.83] =$$

$$= 1 - P[Z < 0.83] = 1 - 0.7967 = 0.2033$$



7.4.3 Distribución de la proporción muestral para muestras grandes

Una importante consecuencia del teorema del límite central es la que atañe a la distribución de la proporción muestral.

En muchos casos el interés del estudio puede estar orientado al comportamiento de una proporción, como sucede cuando nos interesamos por la proporción de artículos defectuosos que pueden darse en un proceso de producción o por la proporción de personas que consumen determinado producto.

Teorema 7.5 Sea $\overline{p}$ la proporción muestral asociada a una característica, la cual se presenta en la población en una proporción π. Entonces, $\overline{p} \sim N\left(\pi, \frac{\pi(1-\pi)}{n}\right)$. Esto es, $Z = \dfrac{\overline{p} - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \sim N(0, 1)$.

Se sabe que la proporción de artículos defectuosos en un proceso de manufactura es de 0.10. El proceso se vigila periódicamente al tomar muestras aleatorias de tamaño 100 e inspeccionar las unidades. Calcule la probabilidad de que esta muestra arroje una proporción de defectuosos a) mayor de 0.17, b) menor de 0.05.

Sea $\overline{p}$ = Proporción de artículos defectuosos hallados en la muestra, $\pi = 0.1$

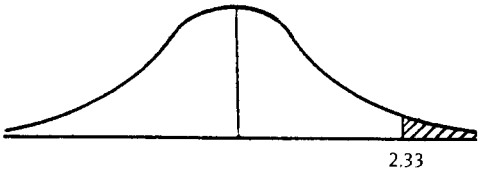
De acuerdo con el teorema 7.5, $\overline{p} \sim N(0.1, 0.1 \times 0.9/100)$.

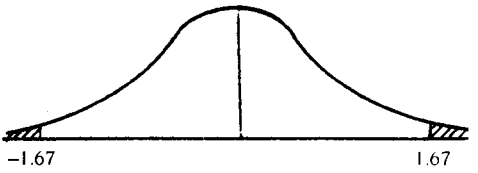
Así que $P[\overline{p} > 0.17] = P\left[\frac{\overline{p} - 0.1}{0.03} > \frac{0.17 - 0.1}{0.03}\right] = P[Z > 2.33]$

$$= 1 - P[Z \leq 2.33] = 1 - 0.9901 = 0.0099$$

Igualmente se tiene $P[\bar{p} < 0.05] = P[\frac{\bar{p} - 0.1}{0.03} < \frac{0.05 - 0.1}{0.03}] = P[Z < 1.67] =$

$= 1 - 0.9525 = 0.0475.$

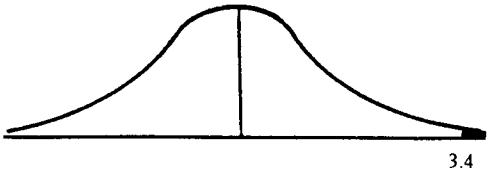




Otro ejemplo. El administrador de la cafetería central de una universidad afirma que sólo el 18% de los estudiantes consumen bebidas calientes. La dirección de la universidad, con el propósito de reunir evidencias para comprobar lo asegurado por el administrador, condujo una encuesta y encontró que de 120 estudiantes seleccionados al azar 36 consumían bebidas calientes en la cafetería. ¿Qué se puede decir, de acuerdo con estos datos sobre lo que dice el administrador?

Para la solución de este problema hacemos un análisis similar al caso de la media. Admitido el hecho de que el administrador está en lo cierto entonces $\bar{p} \sim N(0.18, \frac{0.18 \times 0.82}{120})$ en donde $\bar{p}$ = proporción de estudiantes que consumen bebidas calientes.

La proporción muestral de los estudiantes que consumen bebidas calientes es 36/120 = 0.3.



$$P[\bar{p} > 0.3] = P[\frac{\bar{p} - 0.18}{\sqrt{0.18 \times 0.82/120}} > \frac{0.3 - 0.18}{0.036}] = P[Z > 3.4] = 0.$$

Por tanto, la proporción afirmada por el administrador no parece ser correcta.

Para que la aproximación normal arroje buenos resultados en la solución de problemas relativos a la proporción, además de permanecer constante la probabilidad de éxito, se deben tomar muestras de tamaño grande. Cuando se trate de muestras pequeñas se empleará la distribución binomial o la hipergeométrica.

Ejercicios 7.2

1. Si X tiene distribución normal, con media 10 y varianza 16, y si Y tiene también distribución normal con media 10 y varianza 15. Suponga que X y Y son independientes y determine la distribución de $W = 2X - 3Y$.

2. Suponga que el contenido de nicotina de cierta marca de cigarrillos tiene distribución normal con media $\mu = 25$ miligramos y desviación estándar $\sigma = 4$ miligramos. Se toma una muestra aleatoria de 25 cigarrillos que arroja una media muestral de $\bar{x} = 26$ miligramos. Halle la probabilidad de que una muestra de este tamaño arroje una media muestral de la magnitud dada o mayor.

3. Sea X la velocidad de las mecanógrafas en una gran compañía. Suponga que X tiene distribución normal con media $\mu = 58$ palabras y desviación estándar $\sigma = 16$ palabras. Si se toman 16 mecanógrafas y se les somete a una prueba de velocidad, ¿cuál es la probabilidad de que la media muestral $\bar{X}$ esté entre 50 y 70?

4. Sea X el tiempo de vida de una bombilla marca I, Y la de una de marca II, al ser X y Y independientes. Se sabe que X tiene distribución normal de media $\mu_x = 1{,}000$ horas y varianza $\sigma^2_x = 2{,}500$ horas y que Y tiene distribución normal con media $\mu_Y = 900$ horas y varianza $\sigma^2_Y = 2{,}400$ horas. Sea $D = X - Y$.
a) ¿Cuál es la distribución de la diferencia $X - Y$?
b) ¿Cuál es la media y la desviación estándar de D?

5. Un fabricante de lámparas asegura que la vida promedio de las lámparas que produce es de 1,000 horas con una desviación estándar de 100 horas. Un comprador potencial decide probar si la vida promedio es como lo garantiza el fabricante y para ello toma una muestra de 64 lámparas, las prueba y obtiene una vida media muestral de $\bar{X} = 957$ horas. ¿Qué puede decirse acerca de la vida promedio garantizada por este fabricante?

6. El gerente de una planta asegura que el peso promedio de las cajas del producto que empaca es de 800 gramos con una desviación estándar de 5 gramos. Un inspector seleccionó aleatoriamente 100 cajas y halló un peso promedio de $\bar{x} = 795$ gramos, ¿qué puede decirse del peso promedio que asegura el gerente?

7. Se recibe un lote muy grande de artículos proveniente de un fabricante, el cual asegura que el porcentaje de artículos defectuosos es del 2%. Al seleccionar una muestra aleatoria de 200 artículos y después de inspeccionarlos, se descubren 8 defectuosos. ¿De acuerdo con este valor observado, podemos decir que el fabricante está en lo cierto?

8. En una gran ciudad la proporción de personas que padecen problemas pulmonares debido a la polución, es del 30%. Se escogen 100 personas al azar; halle la probabilidad de que la proporción de los que tengan problemas pulmonares motivados por la polución sea, a) por lo menos del 38%; b) no sea superior al 20%.

9. El jefe de registros de una universidad asegura que el tiempo promedio que requieren los estudiantes para su inscripción de asignaturas es de 30 minutos con una desviación estándar de 5 minutos. Se hizo la anotación del tiempo que necesitaron 64 estudiantes para registrar sus asignaturas y el promedio fue de 34.5 minutos, ¿qué se puede decir del tiempo promedio que ha dado el jefe de registros?

10. El candidato A asegura que el 70% de los votantes lo favorecerán con su voto. Una muestra de 500 electores potenciales dio como resultado que el 65% lo favorecerán con su voto, ¿puede estar tranquilo el candidato A, de acuerdo con esta respuesta del electorado?

7.5 DISTRIBUCIÓN DE LA VARIANZA MUESTRAL. DISTRIBUCIÓN JI CUADRADO

A veces lo que puede interesar en la investigación es la variabilidad de las medidas; por ejemplo, determinar entre dos dispensadores de gaseosa cuál presenta mayor variabilidad en la cantidad de líquido vertido.

La variabilidad, por lo general, se mide mediante la varianza (desviación estándar) y la estadística empleada para llevar a cabo procesos inferenciales es la

varianza muestral que se denota y define como se indicó, de la manera siguiente $S^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2/(n-1)$. El proceso inferencial es posible de llevar a cabo si conocemos el comportamiento probabilístico de la estadística considerada, S^2 en este caso. Tal es el tema del que nos ocuparemos en seguida.

Se dice que una variable aleatoria X tiene distribución ji cuadrado con k grados de libertad, $X \sim \chi^2(k)$ cuando su función de densidad está dada por la fórmula

$$f_x(x) = \begin{cases} \dfrac{1}{\Gamma(k/2)} (1/2)x^{k/2-1}e^{-(1/2)x} & x > 0 \\ 0 \text{ en cualquier otro caso} \end{cases}$$

$\Gamma(k/2)$ es el valor de la función gamma evaluada en $k/2$. Esta función es de gran importancia en los estudios de muchos problemas de la matemática aplicada. χ es una letra griega que se lee ji, chi, o ki. Como ocurre en el caso de variables aleatorias continuas, y la que estamos tratando lo es, los valores de probabilidad de los eventos asociados a la variable se obtienen por integrales definidas (área bajo la curva) de la función de densidad; y como sucede con la mayoría de las variables de este tipo, debido a la complejidad en el cálculo de las integrales, es necesario recurrir a técnicas muy elaboradas para obtener los valores de áreas que luego se registran en una tabla para su posterior uso. Antes de entrar en los detalles del manejo de la tabla de la distribución ji cuadrado, precisemos algunas propiedades de esta distribución.

1. Si X es una variable con distribución ji cuadrado con k grados de libertad $\Longrightarrow$ $E(X) = k$, $V(X) = 2k$.
2. Una variable con distribución ji cuadrado no toma valores negativos.
3. La gráfica de una distribución ji cuadrado es del tipo de curvas sesgadas a la derecha, como se muestra en la figura 7.9.

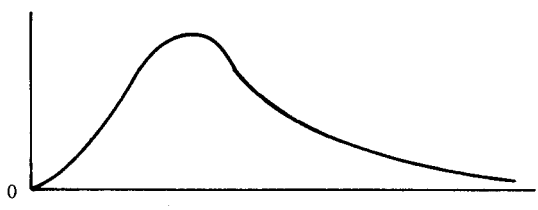


Figura 7.9 *Distribución ji cuadrado.*

4. A medida que aumentan los grados de libertad la curva se va haciendo más simétrica y su cola derecha se va extendiendo, como se muestra en la figura 7.10. De esta figura también queda claro que por cada valor de k se obtiene una distribución distinta; k (grados de libertad) es el único parámetro asociado a la distribución.

Como ya fue insinuado, el área bajo la curva χ^2 representa probabilidades. Puesto que para cada valor de k (grados de libertad) se da una curva distinta, puede pensarse que debería tenerse una tabla para cada valor de k. Sin embargo, debido a que en el proceso inferencial sólo algunos valores de probabilidad (área bajo la curva) resultan pertinentes, se ha convenido tomar en cuenta únicamente esos valores y de esta

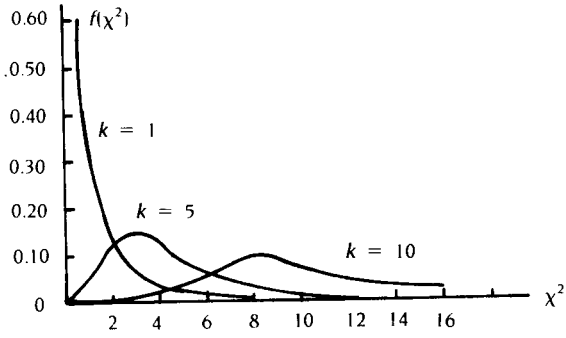
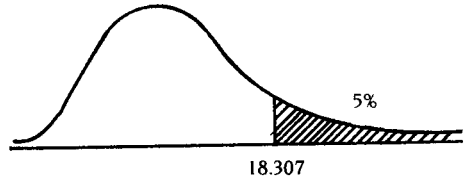


Figura 7.10 *Distribuciones ji cuadrado para distintos grados de libertad.*

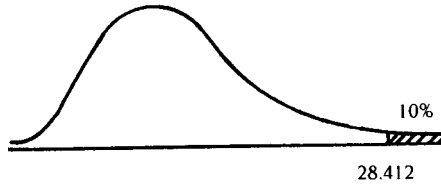
manera la información que nos interesa de la distribución ji cuadrado queda resumida en una tabla sencilla, como la que se da en el apéndice III (Tabla III).

La tabla III da los valores χ^2 que determinan la cola superior (derecha) de la distribución especificada por el número de los grados de libertad. (*Véase* la figura de la parte superior de la tabla). En la fila superior da los valores de cola derecha y en la primera columna encontramos los respectivos grados de libertad. El número que se encuentra en la intersección de la horizontal imaginaria que parte del número dado por los grados de libertad con la vertical, también imaginaria, que sale del valor de la cola derecha es el número que determina dicha cola con los grados de libertad dados. Se denota $\chi^2_{(k,\,q)}$ en donde k son los grados de libertad y q la medida de la cola derecha.

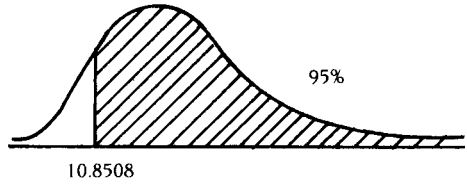
Suponga por ejemplo, que se trata de una variable con distribución ji cuadrado con 10 grados de libertad, $\chi^2(10)$. Se encuentra en la tabla que el valor χ^2 en la columna encabezada con 0.05 con 10 grados de libertad es 18.307. Esto es, que la probabilidad de que una variable arroje un valor χ^2 de 18.307 o superior, es 0.05. Se escribe, $\chi^2_{(10;\,0.05)} = 18.307$ y se representa gráficamente como sigue:



En la distribución ji cuadrado con 20 grados de libertad, ocurrirá un valor χ^2 mayor o igual a 28.412 con una probabilidad de 0.1, se escribe $\chi^2_{(20;\,0.1)} = 28.412$ y se representa gráficamente como sigue:



Si lo que deseamos es determinar valores ji cuadrado que definan regiones al extremo izquierdo de la distribución (cola izquierda), debemos tener en cuenta que una región a la izquierda determina otra a la derecha (no simétrica en este caso) y tenemos así que la cola izquierda del 5% (0.05) corresponde a una región al lado derecho del 95% y así $\chi^2_{(20;\ 0.95)} = 10.8508$ es el número que determina una cola izquierda del 5%. *Véase* figura.



Para valores de k superior a 100, el valor χ^2 se puede aproximar al utilizar la fórmula,

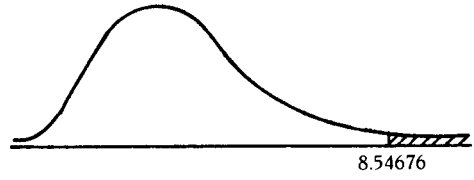
$$\chi^2_{(k;\ q)} = 1/2\left[z_q + \sqrt{2k - 1}\right]^2$$

donde z_q es el valor de la normal estándar correspondiente a una cola derecha de tamaño q. Por ejemplo, si se trata de una distribución χ^2 con 113 grados de libertad, el valor de χ^2 que separa al 5% superior de la curva ($q = 0.05$) es

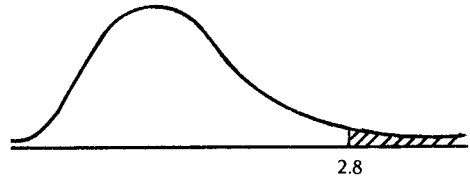
$$\chi^2_{(113;\ 0.05)} = 1/2\left[1.645 + \sqrt{2(113) - 1}\right]^2 = 138.5$$

En algunos casos lo que se necesita es determinar qué porción de cola derecha (izquierda) queda determinada por cierto valor χ^2 con ciertos grados de libertad. Así por ejemplo, halle la porción q de cola derecha para la cual se tiene un valor ji cuadrado con 15 grados de libertad de 8.54676. Esto es, $\chi^2_{(15;\ q)} = 8.54676$.

Para hallar este valor q, nos situamos en el número 15 de los grados de libertad y al movernos horizontalmente hacia la derecha a partir de 15, localizamos el número 8.54676 (o el más próximo); el número que se halle en la parte superior de la vertical que parte de 8.54676 es el valor de q. En el presente caso es $q = 0.9$, como se ilustra en la figura.



Obtenga el valor q que satisface $\chi^2_{(8;\ q)} = 2.8$.
Al buscar en la tabla no lo encontramos exactamente y por ello tomamos el más próximo, 2.7 en este caso, y así $q = 0.95$. *Véase* figura.



Una vez que hemos explicado el manejo de la tabla para una distribución χ^2, entramos a exponer cómo se emplea esta distribución para hacer inferencias relacionadas con la varianza y para una mejor comprensión en su uso explicaremos cómo se origina esta distribución.

Hemos visto que si $X \sim N(\mu, \sigma^2)$, entonces la variable $Z = \frac{X - \mu}{\sigma}$ tiene distribución normal estándar. Pero, ¿qué sucede si tomamos Z^2? Esto es, $Z^2 = \frac{(X - \mu)^2}{\sigma^2}$. La respuesta es que la nueva variable tiene distribución χ^2 con un grado de libertad.

Por otra parte, es un hecho que se demuestra a otro nivel de la estadística el que si se tienen n variables aleatorias independientes $x_1, x_2, \ldots, x_n$ con distribución ji cuadrado con grados de libertad respectivos $k_1, k_2, \ldots, k_n$ entonces la variable $Y = x_1 + x_2 + \ldots + x_n$ tiene distribución ji cuadrado con $k = k_1 + k_2 + \ldots + k_n$ grados de libertad.

Al tener este resultado en mente, podemos afirmar que si $X_1, X_2, \ldots, X_n$ es una muestra aleatoria proveniente de una población de media μ y varianza σ^2, entonces la variable $\chi^2_{(n)} = \frac{(X_1 - \mu)^2}{\sigma^2} + \frac{(X_2 - \mu)^2}{\sigma^2} + \ldots + \frac{(X_n - \mu)^2}{\sigma^2}$ (7-11) tiene una distribución ji cuadrado con n grados de libertad.

El anterior resultado, aunque importante, no es apropiado para hacer inferencias, puesto que en su expresión incluye la media poblacional μ y ésta por lo general también es desconocida.

Así que la expresión (7-11) la sustituimos por $\sum_{i=1}^{n} (X_i - \overline{X})^2/\sigma^2$.

Aritméticamente μ fue sustituido por $\overline{X}$, pero este cambio trae como consecuencia una alteración en los grados de libertad, como se verá a continuación:

$$\sum_{i=1}^{n}(X_i - \mu)^2 = \sum_{i=1}^{n}[(X_i - \overline{X}) + (\overline{X} - \mu)]^2 = \sum_{i=1}^{n}[(X_i - \overline{X})^2 + 2(X_i - \overline{X})(\overline{X} - \mu) + (\overline{X} - \mu)^2] =$$

$$\sum_{i=1}^{n}(X_i - \overline{X})^2 + \sum_{i=1}^{n}(\overline{X} - \mu)^2 + 2\sum_{i=1}^{n}(X_i - \overline{X})(\overline{X} - \mu)$$

Como $(\overline{X} - \mu)$ es una constante y $\sum_{i=1}^{n}(X_i - \overline{X}) = 0 \Longrightarrow 2\sum_{i=1}^{n}(X_i - \overline{X})(\overline{X} - \mu) =$

$(\overline{X} - \mu)\sum_{i=1}^{n}(X_i - \overline{X}) = 0$ y así la anterior expresión se reduce a:

$$\sum_{i=1}^{n}(X_i - \mu)^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2 + \sum_{i=1}^{n}(\overline{X} - \mu)^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2 + n(\overline{X} - \mu)^2$$ Al dividir

entre σ^2 se tiene: $$\frac{\sum_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2} + \frac{n(\overline{X} - \mu)^2}{\sigma^2}$$ (7-12)

Recordemos ahora que cuando X tiene una distribución normal de media μ y varianza σ^2, entonces $\overline{X} \sim N(\mu, \sigma^2/n)$. Así que la variable $\frac{n(\overline{X} - \mu)^2}{\sigma^2} = \frac{(\overline{X} - \mu)^2}{(\sigma^2/n)}$ tiene una distribución ji cuadrado con 1 grado de libertad. Esto es, $\chi^2(1)$.

Además como la varianza muestral $S^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1) \Rightarrow \dfrac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2} = \dfrac{(n-1)S^2}{\sigma^2}$

Por lo anterior, (7-12) puede escribirse como, $\chi^2_{(n)} = \dfrac{(n-1)S^2}{\sigma^2} + \chi^2_{(1)}$ y así $X = \dfrac{(n-1)S^2}{\sigma^2}$ es una variable con distribución ji cuadrado con (n − 1) grados de libertad. Esto es,

$$\chi^2_{(n-1)} = \frac{(n-1)S^2}{\sigma^2} \qquad (7\text{-}13).$$

Lo anterior nos está indicando que al sustituir μ por $\bar{X}$ se pierde un grado de libertad. También es de resaltar que la expresión (7-13) no incluye la media y sólo incluye a σ^2 como parámetro y así mediante ella podemos realizar inferencias respecto de la varianza.

Cuando se trata de probabilidades (inferencia) relacionadas con la varianza, se toma como elemento básico para el análisis la estadística S^2 y mediante las operaciones aritméticas del caso la transformamos en la expresión (7-13) que nos permitirá utilizar la tabla III. Veamos un ejemplo.

Se sabe que la distribución del espesor de cierto material está normalmente distribuida con desviación estándar 0.01 cm. Una muestra aleatoria de 25 piezas de este material arroja como resultado una desviación muestral de 0.008. Halle la probabilidad de observar un valor muestral como éste u otro menor.

Si denotamos la desviación estándar por S, el problema nos solicita $P[S \leq 0.008]$. A continuación se realizan las operaciones del caso hasta llegar a la expresión (7-13).

$$P[S \leq 0.008] = P\left[\frac{(24)S^2}{(0.01)^2} \leq \frac{(24)S^2}{(0.01)^2}\right] = P[X \leq 15.36]$$

Ahora bien, como X tiene la forma de la expresión (7-13), entonces X tiene distribución χ^2 con 24 grados de libertad y así $P[X \leq 15.36] = 0.10$ (aproximadamente).

Finalmente, respecto de la distribución ji cuadrado son pertinentes dos observaciones:

En primer lugar, para que podamos considerar que la variable $\dfrac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n}(X_i - \bar{X})^2/\sigma^2$ tiene distribución ji cuadrado se requiere que las variables $X_1, X_2, \ldots, X_n$ tengan distribución normal y sean independientes.

El segundo aspecto se refiere a los grados de libertad. Esta noción surge del hecho de que si fijamos la ecuación $\sum_{i=1}^{n}(X_i - \bar{X}) = 0$ entonces $(n-1)$ sumandos en esta expresión pueden tomar cualquier valor y el restante tomará el valor del caso para hacer válida la ecuación.

Ejercicios 7.3

1. Calcule los siguientes valores χ^2. $(P(x > x_0))$

a) $\chi^2_{(5,0.95)}$ *b)* $\chi^2_{(30,0.05)}$

c) $\chi^2_{(100,0.90)}$ *d)* $\chi^2_{(120,0.90)}$

2. *a)* Si $X \sim \chi^2(12)$, halle $P[X \leq 6.30]$ *b)* Si $X \sim \chi^2(25)$, halle $P[X > 29.3]$.

3. Si $X \sim \chi^2(3)$ y $Y \sim \chi^2(5)$ e independientes, halle:

a) $P[(X + Y) \leq 13.36]$ *b)* $P[(X + Y) > 15.51]$

4. Si $Z_i \sim N(0,1)$, calcule $P[\sum_{i=1}^{n} Z_i^2 \leq 4.86]$

5. Se obtiene una muestra aleatoria de tamaño $n = 16$ de una distribución normal con media y varianza desconocidas, obtenga

$$P[\frac{S^2}{\sigma^2} \leq 2.04]$$

6. En la producción de cierto material para soldar se ha establecido que la desviación estándar de la tensión a la ruptura de este material es de 25 libras. Una muestra de tamaño 36 dio una desviación estándar de 25.8 libras. ¿Cuál es la probabilidad de observar un valor muestral u otro mayor?

7. Halle el valor q para el cual se tiene:

a) $\chi^2_{(5,q)} = 11.0705$ *b)* $\chi^2_{(12,q)} = 4.40379$

7.6 DISTRIBUCIÓN *t* DE STUDENT

Hemos indicado que si se tiene una muestra aleatoria $X_1, X_2, \ldots, X_n$ proveniente de una población normal de media μ y varianza σ^2 entonces la variable

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

Para poder emplear esta variable con propósitos inferenciales respecto de la media, es necesario conocer σ; sin embargo, lo más común es que σ también sea desconocida y por ello debe ser estimada. Si en la anterior expresión para Z reemplazamos σ por S (desviación muestral) obtenemos la expresión $T = \frac{\sqrt{n}(\overline{X} - \mu)}{S}$, y de esta variable se dice que tiene distribución t de Student con $(n - 1)$ grados de libertad.

Formalmente una variable T con distribución t de Student o simplemente distribución t se define de la forma siguiente:

Se dice que una variable T tiene distribución t de Student con k grados de libertad, $T \sim t(k)$, si su función de densidad está dada por

$$f_T(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \left[1 + \frac{t^2}{n}\right]^{-(n+1)/2}, \; t \in R. \; \Gamma\left(\frac{n+1}{2}\right), \Gamma\left(\frac{n}{2}\right) \text{ son valores}$$

de la función gamma.

Al ser la distribución t del tipo continuo, sus valores de probabilidad también corresponden a áreas bajo la curva y sus valores se obtienen mediante consultas de tablas construidas para tal fin.

Algunas propiedades de la distribución t son:

1. Si T es una variable aleatoria con distribución t con k grados de libertad $\Longrightarrow E(T) = 0$ y $V(T) = \frac{n}{n-2}$, $n > 2$.
2. La variable T teóricamente toma valores desde $-\infty$ a ∞.

3. La gráfica de $f(t)$ es de forma de campana, como se indica en la figura 7.11.

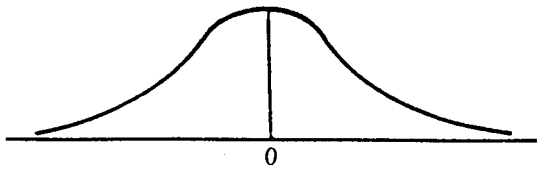


Figura 7.11 *Distribución t de Student.*

4. La gráfica de la distribución t es parecida a la de la normal estándar y se diferencia fundamentalmente en que:

i) En las colas la t está por encima de la normal estándar
ii) En el centro la t está por debajo de la normal estándar. *Véase* figura 7.12.

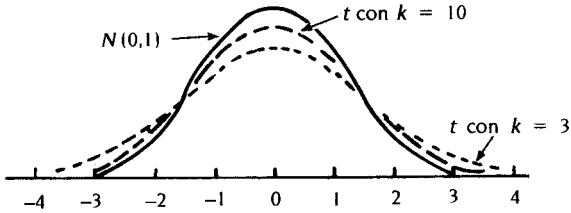


Figura 7.12 *La distribución t comparada con la distribución normal estándar.*

5. Cuando los grados de libertad son altos, los valores t coinciden con los de la normal estándar.

Al igual que ocurre con la distribución χ^2, cada valor de grado de libertad determina una distribución t distinta, pero debido a que sólo algunos valores de área son de importancia, la información pertinente para la t se reduce a una tabla. Y esto es lo que ocurre con la tabla IV.

Para obtener valores de área (probabilidad) de la distribución t al emplear la tabla IV tenemos en cuenta que cada una de las probabilidades que aparece en la segunda fila es igual a la suma de las áreas de las dos colas, según se indica en el diagrama de la parte superior de la tabla, en tanto que las que aparecen en la fila Q, cada una es igual a una de las áreas de las dos colas de una distribución dada. Para cada uno de los valores de k que aparecen en la columna izquierda de la tabla, la partida de la tabla en una columna particular es el valor t por encima del cual queda la probabilidad que aparece en la fila Q. Si bien sólo aparecen en la tabla los valores positivos de t, quedan implícitos los valores negativos, ya que la distribución es simétrica respecto de 0. Así, para 10 grados de libertad ($n = 11$), la probabilidad es 0.10 de que los valores t sean mayores que 1.372, valor que se encuentra en la intersección de la fila 10 y la columna 0.1 para Q (0.2 para $2Q$) y escribimos $t_{(10,0.1)} = 1.372$ y se representa gráficamente como se muestra en la figura 7.13.

En resumen, cuando se trata de inferencia respecto de la media en poblaciones normales de varianza desconocida, emplearemos la variable $T = \frac{\sqrt{n}(\overline{X} - \mu)}{S}$ la cual tiene una distribución t con $(n - 1)$ grados de libertad.

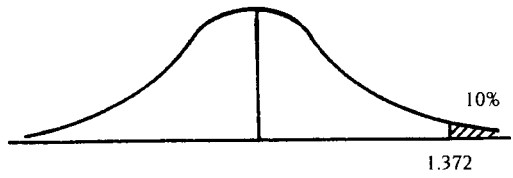


Figura 7.13 *El área sombreada corresponde al 10% del área total.*

En el siguiente ejercicio se ilustra cómo se utiliza la distribución t en la solución de problemas:

El gerente de una fábrica de cierto tipo de alimentos asegura que el peso promedio del producto que elabora es de 165 g. Un consumidor desconfiado para probar lo afirmado por el gerente decide escoger 16 paquetes del producto y pesarlos. Los resultados fueron los siguientes: 165, 158, 153, 162, 171, 175, 173, 169, 166, 170, 164, 177, 148, 167, 152, 149. ¿Evidencian estos datos que el gerente está en lo cierto? Suponga que los pesos se distribuyen normalmente.

La solución a este problema la encontramos al seguir un procedimiento similar al que empleamos antes. Así que partimos de que es correcto lo planteado por el gerente. Por tanto, la distribución de los pesos es como se muestra en la figura 7.14.

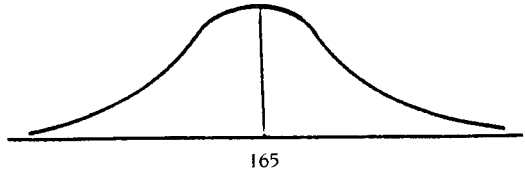
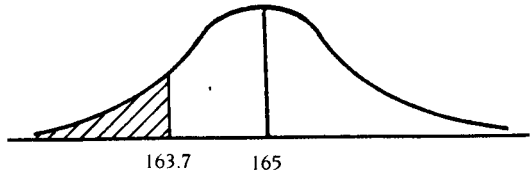


Figura 7.14 *Distribución de los pesos del producto.*

Ahora ubicamos la media muestral en el eje horizontal.

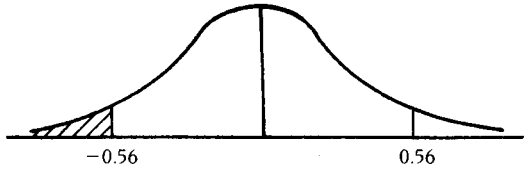
A partir de los datos obtenemos que la media y la desviación muestral están dadas por $\bar{x} = 163.6875$, $S = \sigma_{n-1} = 9.24$



Debemos calcular $P[\overline{X} \leq 163.7]$. Como se trata de población normal y varianza desconocida, empleamos la distribución t.

$$P\left[\frac{\sqrt{16}(\overline{X} - 165)}{9.24} \leq \frac{\sqrt{16}(163.7 - 165)}{9.24}\right] = P[T \leq -0.56] = P[T \geq 0.56]$$

que corresponde a una probabilidad mayor del 40%.



Por tanto, no se puede rechazar lo garantizado por el gerente.

Ejercicios 7.4

1. Resuma las características de la distribución t.
2. Explique en qué se diferencian la distribución t y la normal.
3. Determine los siguientes valores e ilústrelos gráficamente.

   a) $t_{(5,\ 0.9)}$ *b)* $t_{(12,\ 0.975)}$ *c)* $t_{(3,\ 0.05)}$

4. Dada una muestra aleatoria de n observaciones $X_1, X_2, \ldots, X_n$ donde $X \sim N(5, \sigma^2)$ y que da un cociente $\frac{\sqrt{n}(\bar{X} - \mu)}{S}$ de distribución t con $(n - 1)$ grados de libertad, responda lo siguiente:

   a) Si $n = 25$, $\bar{x} = 3$ y $S = 2$, ¿cuál es el valor de t?
   b) Si $n = 9$, $\bar{x} = 2$ y $t = -2$, ¿cuál es el valor de S?
   c) Si $n = 25$, $S = 10$ y $t = 2$, ¿cuál es el valor de $\bar{x}$?
   d) Si $S = 15$, $\bar{x} = 14$ y $t = 3$, ¿cuál es el valor de n?

5. Halle el valor q tal que:

   a) $t_{(5,\ q)} = 2.571$ *b)* $t_{(19,\ q)} = 0.688$

6. Los siguientes son los tiempos que tardan 6 trabajadores de una gran empresa en tomar el almuerzo: 27, 15, 20, 32, 18 y 26 minutos. ¿Justifican estos datos la afirmación según la cual el tiempo promedio que tardan los empleados en almorzar es de 20 minutos? Suponga que los tiempos se distribuyen normalmente.

7.7 DISTRIBUCIÓN DE LA DIFERENCIA DE MEDIAS EN POBLACIONES NORMALES INDEPENDIENTES

Es frecuente interesarse por la *diferencia* entre dos medias; por ejemplo, comparar el contenido promedio por botella que proviene de dos embotelladoras.

La comparación entre medias de dos poblaciones se realiza mediante la variable $D = \bar{X} - \bar{Y}$.

Si tanto X como Y están distribuidas normalmente y además son independientes, la distribución de $D = \bar{X} - \bar{Y}$ puede determinarse según el caso de la forma como se señala en el siguiente teorema:

Teorema 7.6 Suponga que de una población normal X de media μ_x y varianza σ^2_x se extraen muestras de tamaño n_1; de una población también normal Y de media μ_Y y varianza σ^2_Y se extraen muestras de tamaño n_2. Si X y Y son independientes $\Longrightarrow$

$(\bar{X} - \bar{Y}) \sim N(\mu_x - \mu_Y, \frac{\sigma_x^2}{n_1} + \frac{\sigma^2}{n_2})$. Esto es, $Z = \dfrac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_Y)}{\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_Y^2}{n_2}}}$

tiene distribución normal estándar.

Si de la población normal X con media $\mu_x = 106$ y varianza $\sigma^2_x = 240$ y de una población normal Y (independiente) con media $\mu_Y = 95$ y varianza $\sigma^2_Y = 350$ se extraen muestras de tamaños $n_1 = 40$ y $n_2 = 35$ respectivamente. Calcule a) $P[(\bar{X} - \bar{Y}) > 18]$, b) $P[8 < (\bar{X} - \bar{Y}) < 20]$.

Con base en el teorema 7.6 $(\overline{X} - \overline{Y}) \sim N(106 - 95, 240/40 + 350/35)$. Esto es, $(\overline{X} - \overline{Y}) \sim N(11, 16)$. Por tanto,

$$a)\ P[(\overline{X} - \overline{Y}) > 18] = P[\frac{(\overline{X} - \overline{Y}) - 11}{4} > \frac{18 - 11}{4}] = P[Z > 1.75] =$$

$$= 1 - P[Z \leq 1.75] = 1 - 0.9599 = 0.0401$$

$$b)\ P[8 < (\overline{X} - \overline{Y}) < 20] = P[\frac{8 - 11}{4} < \frac{(\overline{X} - \overline{Y})}{4} < \frac{20 - 11}{4}] =$$

$$= P[-0.75 < Z < 2.25] = 0.9878 - 0.2266 = 0.7612$$

Para que sea posible la obtención de probabilidades de acuerdo con el teorema 7.6, es necesario conocer las varianzas de ambas poblaciones. Sin embargo, lo más frecuente es que tales varianzas sean desconocidas, pero son consideradas iguales. Cuando esto ocurre, la diferencia de las medias se analiza al utilizar la variable T que se define en el siguiente teorema:

Teorema 7.7 Suponga que de una población normal X con media μ_x y varianza σ^2_x desconocida se extrae una muestra de tamaño n_1; de una población normal Y con media μ_Y y varianza σ^2_Y desconocida se extraen muestras de tamaño n_2. Si X y Y son independientes y $\sigma^2_x = \sigma^2_Y \Longrightarrow$ La variable $T = \dfrac{(\overline{X} - \overline{Y}) - (\mu_x - \mu_Y)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ tiene una distribución t con $(n_1 + n_2 - 2)$ grados de libertad.

S_p es la raíz cuadrada de $S^2_p = \dfrac{(n_1 - 1)S^2_1 + (n_2 - 1)S^2_2}{(n_1 + n_2 - 2)}$ en donde S^2_1 y S^2_2 son las respectivas varianzas muestrales. S^2_p se llama varianza ponderada.

El siguiente ejercicio nos permitirá comprender mejor el uso de la distribución t en el estudio de diferencia de medias.

Dos grupos de trabajadores se sometieron a una prueba consistente en la medición del tiempo que necesitó cada uno de ellos para llevar a cabo una labor específica. Los tiempos (en minutos) fueron como se anota en la siguiente tabla:

Grupo I	Grupo II
15.3	21.2
18.7	22.4
22.3	18.3
17.6	19.3
19.1	17.1
14.8	27.7

a) Halle la diferencia del tiempo promedio muestral del grupo II y el grupo I.
b) Suponga que el tiempo promedio requerido por los dos grupos es igual y halle la probabilidad de obtener un promedio de diferencia mayor o igual a $\overline{y}$–$\overline{x}$.

Definimos las variables,
X = Tiempo que gasta un individuo del grupo I en realizar la tarea
Y = Tiempo que gasta un individuo del grupo II en realizar la tarea

Calculamos el tiempo promedio y la desviación estándar muestral para el grupo I, $\bar{x} = 17.97$, $S = \sigma_{n-1} = 2.75$

Igualmente, se tiene el tiempo promedio y la desviación estándar para los tiempos del grupo II, $\bar{y} = 21$, $S = \sigma_{n-1} = 3.8$

De lo anterior se tiene $\bar{y} - \bar{x} = 3.03, S^2_p = \dfrac{5(2.75)^2 + 5(3.8)^2}{6 + 6 - 2} = 11 \; (S_p = 3.32)$

Ahora calculamos $P[(\bar{Y} - \bar{X}) \geq 3.03] = P\left[\dfrac{(\bar{Y} - \bar{X})}{(3.32)\sqrt{1/6 + 1/6}} \geq \dfrac{3.03}{(3.32)\sqrt{1/6 + 1/6}}\right]$

$= P[T \geq 1.58] = 0.075$ (al hacer una interpolación).

7.8 DISTRIBUCIÓN DEL COCIENTE DE DOS VARIANZAS. DISTRIBUCIÓN *F*

En la sección anterior se estudió la diferencia de medias. Pero, a menudo se encuentra la situación en que se requiere la comparación entre dos varianzas de población; es decir, determinar si la variabilidad de una población difiere de otra. La distribución *F* se emplea para estos casos.

La *distribución F*, así conocida en honor a R. A. Fisher, el primero en desarrollarla y describirla, está relacionada con la estadística (variable aleatoria)

$$F = \frac{S^2_X}{S^2_Y} \tag{7-14}$$

en donde S^2_X y S^2_Y son las varianzas muestrales de las dos poblaciones.

La distribución *F* además de ser un medio que nos permite hacer inferencias respecto de la varianza de dos poblaciones, también nos proporciona un modo para comparar las medias de tres o más poblaciones mediante la técnica conocida como *análisis de varianza*.

Como ocurre con cualquier otra distribución, la distribución *F* está caracterizada por su función de densidad, la cual está definida de la forma siguiente:

$$f_x(x) = \begin{cases} \dfrac{\Gamma(m + n)/2}{\Gamma(m/2)\,\Gamma(n/2)} \left(\dfrac{m}{n}\right)^{m/2} \dfrac{x^{(m-n)/2}}{[1 + (m/n)\,x]^{(m+n)/2}} & x > 0 \\ 0, \text{ en cualquier otro caso} \end{cases}$$

$\Gamma(m + n)/2$, $\Gamma(m/2)$ $\Gamma(n/2)$ son valores de la función gamma.

La complejidad matemática de la expresión que define la función de densidad de una variable aleatoria con distribución *F* no nos permite ver con claridad las características de ésta y sólo lo que podemos decir aquí es que al igual que la distribución ji cuadrado, la distribución *F* no toma valores negativos y tiene asociados grados de libertad que constituyen sus parámetros.

La representación gráfica de la distribución *F* depende de sus grados de libertad. *Véase* figura 7.15.

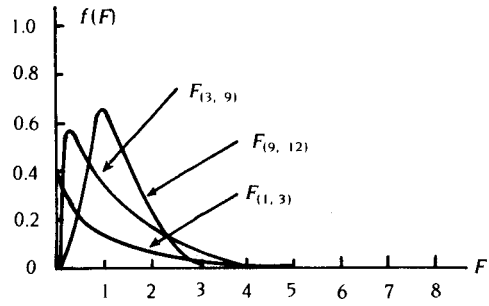


Figura 7.15 *Distribución F para distintos grados de libertad.*

La distribución F, como ya se dijo al comienzo de esta sección, está relacionada con el cociente de varianzas $\frac{S^2_X}{S^2_Y}$ en donde S^2_X y S^2_Y son las varianzas muestrales. La forma como este cociente la determina, es como sigue:
Supongamos que dos variables aleatorias independientes X, Y sean ambas normales, esto es, $X \sim N(\mu_X, \sigma^2_X)$, $Y \sim N(\mu_Y, \sigma^2_Y)$. Si tomamos dos muestras de X y de Y respectivamente de tamaño m y n, y se obtienen las estimaciones S^2_X, S^2_Y de las varianzas poblacionales, entonces como ya se estableció en la sección 7.6,

$$\chi^2_{(m-1)} = \frac{(m-1)S^2_X}{\sigma^2_X} \text{ y } \chi^2_{(n-1)} = \frac{(n-1)S^2_Y}{\sigma^2_Y} \text{ y así al hacer el cociente tenemos,}$$

$$\frac{\chi^2_{(m-1)}/(m-1)}{\chi^2_{(n-1)}/(n-1)} = \frac{[(m-1)S^2_X/\sigma^2_X]/(m-1)}{[(n-1)S^2_Y/\sigma^2_Y]/(n-1)} = \frac{(m-1)S^2_X/(m-1)\sigma^2_X}{(n-1)S^2_Y/(n-1)\sigma^2_Y}$$

$$= \frac{\sigma^2_Y}{\sigma^2_X} \times \frac{S^2_X}{S^2_Y} \Longrightarrow F = \frac{\sigma^2_Y}{\sigma^2_X} \times \frac{S^2_X}{S^2_Y} \qquad (7\text{-}14)$$

La variable F definida por (7-14) se dice que tiene distribución F con $(m-1)$ grados de libertad en el numerador y $(n-1)$ grados de libertad en el denominador y se denota $F \sim F_{(m-1,\, n-1)}$

En cuanto al manejo de la tabla de la distribución F presenta como hecho particular que hay que considerar dos números para los grados de libertad. Esto hace que dicha distribución en cuanto a su tabulación sea bastante extensa comparada con la tabulación de las otras distribuciones que hemos considerado hasta ahora, como la χ^2 y la distribución t. La tabulación de la distribución F se puede hacer de distintas formas; la que se ofrece en la tabla V es una de ellas.

Dado un valor de área q, la fila superior de la tabla enumera los valores n_1, llamados grados de libertad del *numerador*; la primera columna enumera los valores n_2 llamados grados de libertad del *denominador* y la partida de la tabla es el valor F, que se suele designar por $F_{(q;\, n_1;\, n_2)}$ que separa la parte q superior de la distribución F con n_1 y n_2 grados de libertad.

La tabla se ha construido al suponer que la varianza más grande está en el numerador de la ecuación (7-14) y la varianza más pequeña está en el denominador. Por consiguiente, el valor F ha de ser siempre mayor que 1. Solamente cuando m y n tienden ambas hacia infinito.

Vamos a buscar $F_{(0.1;\, 10,\, 15)}$ con el empleo de la tabla V.

Nos remitimos a la parte de la tabla encabezada con la leyenda "valores del 10 por ciento superior". En la horizontal, grados de libertad del denominador, buscamos el número 10 y en la vertical, grados de libertad del denominador, ubicamos el número 15. En la intersección de las dos rectas imaginarias que parten de estos números encontramos el número 2.06. Este número se interpreta de la manera siguiente: "la probabilidad de observar un valor F igual o mayor a 2.06 es de 0.1".
Igualmente tenemos $F_{(0.05;\ 12;\ 20)} = 2.28$
¿Qué debemos hacer si necesitamos $F_{(0.9;\ 15;\ 20)}$, por ejemplo?
Para encontrar valores como este aplicamos la siguiente fórmula válida para la distribución F

$$F_{(q;\ n_1;\ n_2)} = \frac{1}{F_{(1-q;\ n_2;\ n_1)}}$$

En nuestro caso, $F_{(0.9;\ 15;\ 20)} = \dfrac{1}{F_{(0.1;\ 20;\ 15)}} = \dfrac{1}{1.92} = 0.52$

7.9 MUESTREO EN POBLACIONES FINITAS

Como ya lo hemos señalado al comienzo, sólo en poblaciones finitas es posible pensar en la alternativa de muestreo con repetición o muestreo sin repetición. También se ha señalado, y así se ha comprobado, que cuando se hace muestreo con repetición a partir de poblaciones finitas, los resultados se asimilan al caso de muestreo en poblaciones infinitas, en donde por naturaleza el muestreo es con repetición. Cuando se hace dicho muestreo, sea la población finita o infinita, la estadística muestral $\bar{X}$ presenta una media igual a la de la población y varianza la de la población dividida entre el tamaño de la muestra. Esto es, $\mu_{\bar{X}} = \mu_X$, $\sigma^2_{\bar{X}} = \sigma^2{}_X/n$. Pero si el muestreo es sin repetición, ¿se cumple la misma relación?

Para responder este interrogante, tomaremos en consideración la población $S = \{2, 4, 6, 8\}$ que ya se tuvo en cuenta al comienzo del capítulo cuando se estudió el caso del muestreo con repetición. Las muestras posibles de tamaño dos sin repetición que podemos tomar de esta población son las que se dan (también los valores de $\bar{X}$) en la tabla 7.7, en donde X_1 representa el valor de la primera observación y X_2 el valor de la segunda.

Tabla 7.7 *Diferentes muestras posibles de tamaño dos sin repetición tomadas de* $S = \{2, 4, 6, 8\}$.

Muestra	X_1	X_2	$\bar{X}$
1	2	4	3
2	2	6	4
3	2	8	5
4	4	2	3
5	4	6	5
6	4	8	6
7	6	2	4
8	6	4	5
9	6	8	7
10	8	2	5
11	8	4	6
12	8	6	7

A partir de la tabla 7.7 se obtiene la tabla 7.8 que muestra la distribución de $\bar{X}$.

Tabla 7.8

Media muestral (X)	Muestra	Probabilidad, $P(\bar{x})$
3	2	$\frac{2}{12} = \frac{1}{6}$
4	2	$\frac{2}{12} = \frac{1}{6}$
5	4	$\frac{4}{12} = \frac{2}{6}$
6	2	$\frac{2}{12} = \frac{1}{6}$
7	2	$\frac{2}{12} = \frac{1}{6}$

Es claro que como la población permanece fija μ_X y σ^2_X conservan sus valores. Esto es, $\mu_X = 5$ y $\sigma^2_X = 5$

¿Ocurre algún cambio para $\mu_{\bar{x}}$?

$$\mu_{\bar{X}} = 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{2}{6}\right) + 6\left(\frac{1}{6}\right) + 7\left(\frac{1}{6}\right) = \frac{3 + 4 + 10 + 6 + 7}{6}$$

$$= \frac{30}{6} = 5$$

Así que $\mu_{\bar{x}} = \mu_X$, lo mismo que ya se había tenido para el muestreo con repetición.
¿Es la varianza $\sigma^2_{\bar{x}}$ la misma que la del muestreo con repetición?
$\sigma^2_{\bar{x}} = E(\bar{X}^2) - [E(\bar{X})]^2$

$$E(\bar{X}^2) = 3^2\left(\frac{1}{6}\right) + 4^2\left(\frac{1}{6}\right) + 5^2\left(\frac{2}{6}\right) + 6^2\left(\frac{1}{6}\right) + 7^2\left(\frac{1}{6}\right) = \frac{9 + 16 + 50 + 36 + 49}{6} =$$

$$= \frac{160}{6} = \frac{80}{3}$$

$$\sigma^2_{\bar{x}} = \frac{80}{3} - (5)^2 = \frac{5}{3}$$

Al recordar que cuando se consideró el muestreo con repetición $\sigma^2_{\bar{x}} = 5/2$, entonces podemos notar que las varianzas en los dos tipos de muestreo difieren. Esto siempre ocurre en estos muestreos cuando de poblaciones finitas se trata.
¿Cuál es la relación, si existe, entre $\sigma^2_{\bar{x}}$ y σ^2_X?

Cuando el muestreo se hace sin repetición se puede demostrar que $\sigma^2_{\bar{x}} = \frac{N - n}{N - 1} \frac{\sigma_X^2}{n}$ en donde N es el tamaño de la población y n el tamaño de la muestra. Esta relación se puede comprobar fácilmente en el problema que venimos analizando. En efecto, para este caso $N = 4$ y $n = 2$, y así $\frac{N - n}{N - 1} \frac{\sigma^2_X}{n} = \frac{4 - 2}{4 - 1} \frac{5}{2} = \left(\frac{2}{3}\right)\left(\frac{5}{2}\right) = \frac{5}{3}$

En resumen, tenemos respecto del muestreo sin repetición el siguiente teorema:

Teorema 7.8 Sea X una variable aleatoria con media muestral $\bar{X}$ definida en una población finita. Si el muestreo se hace sin repetición $\Longrightarrow$ $\mu_{\bar{X}} = \mu_X$ y $\sigma^2_{\bar{X}} = \left(\frac{N-n}{N-1}\right)\frac{\sigma^2_X}{n}$

Como se puede observar, la varianza de $\bar{X}$ cuando se hace este tipo de muestreo es igual a la varianza de $\bar{X}$ cuando se hace con repetición multiplicada por $\frac{N-n}{N-1}$, este número se llama *factor de corrección para población finita* y sólo se tiene en cuenta cuando el cociente n/N, llamado *fracción de muestreo* es mayor de 0.05. Finalmente, al ser $\frac{N-n}{N-1} < 1$, entonces la varianza de $\bar{X}$ cuando se hace sin repetición es menor que al hacerse con repetición. Esto es, $\bar{X}$ presenta menor variabilidad cuando el muestreo es sin repetición.

Ejercicios 7.5

1. Explique la relación entre las distribuciones F y χ^2.

2. Explique cómo se relaciona la distribución F con la distribución normal.

3. Describa algunas características importantes de la distribución F.

4. Halle los valores k para las probabilidades siguientes:

   a) $P[F \geq k] = 0.05$, dados $m = 12$ y $n = 13$
   b) $P[F \geq k] = 0.01$, dados $m = 5$ y $n = 40$
   c) $P[F \geq k] = 0.025$, dados $m = 40$ y $n = 10$
   d) $P[F \geq k] = 0.95$, dados $m = 10$ y $n = 20$
   e) $P[F \geq k] = 0.99$, dados $m = 20$ y $n = 10$
   f) $P[F \leq k] = 0.01$, dados $m = 10$ y $n = 20$
   g) $P[F \leq k] = 0.05$, dados $m = 20$ y $n = 10$

5. Una población está compuesta por $N = 5$ niños cuyas edades son 2, 3, 5, 6, 8. Se escoge un niño al azar. Considere la variable X, definida como $X =$ edad del niño.

   a) Describa todas las posibles muestras de tamaño dos con repetición.
   b) Describa todas las posibles muestras de tamaño dos sin repetición.
   c) Compruebe que $\mu_{\bar{X}} = \mu_X$, $\sigma^2_{\bar{X}} = \sigma^2_X/n$ cuando el muestreo se hace con repetición.
   d) Compruebe que $\mu_{\bar{X}} = \mu_X$, $\sigma^2_{\bar{X}} = \frac{N-n}{N-1}\frac{\sigma^2_X}{n}$