CAPÍTULO 7

DISTRIBUCIONES EN EL MUESTREO

7.1. Introducción

Generalmente, las poblaciones tienen tamaños que hacen que estudiarla en su totalidad sea poco práctico desde diversos puntos de vista; costo, tiempo, tipo de investigación, etc.. Es necesario por tanto, considerar la selección de una muestra representativa de dicha población y cuyo tamaño sea más manejable.
Si se hace uso de toda la información, se dice que se ha hecho una investigación exhaustiva o total. La alternativa a esta es la investigación parcial. La misma consiste en realizar el análisis en base a la información correspondiente a un subconjunto de los elementos o individuos, una muestra, de forma tal que a un costo y esfuerzo razonable se logren obtener conclusiones tan validas como las que se obtendrían realizando una investigación exhaustiva o total, un censo. Considérese los siguientes ejemplos:

1. Para conocer la nota promedio de los estudiantes de la Universidad de Los Andes (Núcleo Mérida), se debe ir a las oficinas de registros estudiantiles de todas las facultades y solicitar allí las notas de los estudiantes, dicha tarea no es fácil por distintas razones, entre las cuales podemos mencionar la confidencialidad de la información. Por tal razón, a través de una encuesta a cierto número de estudiantes puede determinarse la nota promedio de dicho grupo, y a partir de ese resultado dar una conclusión sobre la población.

2. Si se quisiera conocer el sueldo promedio del venezolano, sería difícil tener acceso al sueldo de todos los venezolanos, al igual que en el caso anterior sólo se podría obtener dicha información de una parte de los venezolanos.

3. Para determinar el nivel de aceptación o rechazo que tiene un candidato a gobernador, no es necesario realizar el sondeo de opinión sobre todos los habitantes del Estado, aún queriendo

recoger dicha opinión sería muy costosa. Es por ello que las empresas encuestadoras realizan el sondeo sobre una parte de la población y a partir de ella interpretan como está el candidato en dicho Estado.

De esta forma, la teoría del muestreo es de gran importancia en estadística, pues permite entre otras cosas, desarrollar: procesos de estimación de características de la población, verificación de afirmaciones o conjeturas (pruebas de hipótesis) que se hacen acerca del verdadero valor de un parámetro, determinar la significación estadística de diferencias que se puedan observar entre poblaciones e identificar relaciones entre variables.
Dada la naturaleza aleatoria de los datos obtenidos en la muestra, hay un riesgo en la certeza de la afirmación propuesta, y es necesario establecer una medida y determinar el valor de este riesgo.
En este capitulo se da una introducción sobre las técnicas de muestreo y se define la distribución muestral de los principales parámetros de una población.

7.2. Muestreo y tipos muestreo

Cuando se habla de investigación, se hace referencia al análisis de características de interés en una población. Este análisis puede hacerse en base a todos los datos o en base a una parte de estos. Como se dijo antes, por lo general no es posible contar con todos los datos, en su lugar es necesario llevar a cabo una investigación parcial, una **investigación por muestreo**, la cual consiste en hacer el análisis en base a una parte o porción del conjunto total de datos. Esta porción recibe el nombre de **Muestra**.
La exactitud de la investigación depende en gran parte de la forma en que la muestra es escogida. La misma debe ser seleccionada de forma tal que sea lo más representativa posible del total y así el riesgo de obtener resultados erróneos sea el mínimo. De aquí que la selección de muestras es un aspecto de crucial importancia en Estadística.
Ahora bien, cuando se selecciona una muestra se deben tomar en cuenta las siguientes consideraciones: Elegir el tamaño de la muestra, lo cual depende no solamente de la cantidad de información que se quiere conseguir, y el grado de certeza deseada, sino también del costo del muestreo y la selección de los elementos que la constituyen. Cualquiera sea el método elegido, el requisito más importante es que la muestra obtenida proporcione una imagen tan real como sea posible de aquella población que sé ha sometido al muestreo.

Definición 7.1 (Muestreo) *Proceso de medición de la información en solo una parte de la población estadística. Se define como el proceso de seleccionar un número de observaciones (sujetos) de un grupo en particular de la población (métodos para seleccionar muestras), que se utiliza cuando no es posible contar o medir todos los elementos de la población objeto de estudio.*

Es práctica común seleccionar una muestra en forma intencional, de acuerdo a opiniones o criterios personales. Una de las principales razones que da origen a esta forma de selección es el de obtener información sin mucho costo. A este tipo de muestreo se le denomina **Muestreo no Probabilístico**.

El mismo, no involucra ningún elemento aleatorio en el procedimiento de selección. Es importante acotar que en condiciones apropiadas estos métodos pueden ofrecer resultados útiles, por ejemplo, cuando solo se necesitan estimaciones toscas las cuales no van a ser utilizadas para tomar decisiones importantes.
Son ejemplos de muestreos no probabilísticos:

1. La muestra es restringida a la parte de la población que es fácilmente accesible.

2. La muestra consiste de los elementos que estén más a la mano

3. Se selecciona un grupo de unidades tipo.

4. La muestra esta compuesta por voluntarios.

Al utilizar métodos no probabilísticos, no todos los miembros de la población tienen probabilidad de ser seleccionados, lo que se traduce en la obtención de una muestra no representativa de la población.
La alternativa ideal es el uso del **Muestreo Probabilístico**. Este procedimiento da a cada elemento de la población una probabilidad de ser seleccionada. De esta forma, existe la posibilidad de que la muestra sea representativa de la población.
Existen métodos de muestreo que se consideran básicos y que dan origen a otros métodos más complejos. Estos son:

1. Muestreo Aleatorio Simple.

2. Muestreo Estratificado Aleatorio Simple.

3. Muestreo Sistemático Aleatorio.

4. Muestreo por Conglomerados.

5. Muestreo Polietápico.

Definición 7.2 (Muestreo Aleatorio Simple) *Denominado también irrestricto aleatorio, es aquel en el que todas las muestras de tamaño n, tienen la misma probabilidad de ser seleccionadas y el método de selección debe verificar que en cualquier fase de la obtención de la muestra cada individuo que no ha sido sacado previamente, tiene la misma probabilidad de ser elegido. Existen dos formas de extraer una muestra de una población: con reposición y sin reposición.*

1. *Muestreo sin reposición. Cuando un elemento ha sido seleccionado y éste no es reemplazado(devuelto a la población), evitando así que la misma unidad entre en la muestra más de una vez, se dice que el muestreo es con reemplazo.*

2. *Muestreo con reposición. Cuando un elemento ha sido seleccionado y éste es reemplazado(devuelto a la población), permitiendo así que la misma unidad entre en la muestra más de una vez, se dice que el muestreo es sin reemplazo. Este tipo de muestreo, es totalmente factible pero, rara vez es usado. Su uso se da básicamente con fines teóricos.*

Definición 7.3 (Muestreo Estratificado Aleatorio Simple) *La población se divide en estratos (sub-poblaciones): luego se toma una muestra aleatoria simple de cada estrato. Los estratos deben ser mutuamente excluyentes y exhaustivos. Los estratos deben ser internamente homogéneos.*

Definición 7.4 (Muestreo Sistemático Aleatorio) *Los elementos de la muestra son seleccionados a intervalos fijos. Se reparte más uniformemente en la población. Es más preciso que el muestreo aleatorio simple.*

Definición 7.5 (Muestreo por Conglomerados) *La población se divide en grupos y luego se elige aleatoriamente una muestra de conglomerados. Los conglomerados son grupos de elementos cercanos desde algún punto de vista (Administrativa o físicamente).*

Definición 7.6 (Muestreo Polietápico) *Diseño muestral que consta de varias etapas en las cuales se usa un diseño simple (diseños que comprenden una sola etapa de clasificación, por ejemplo; muestreo aleatorio simple, etc.).*

7.2.1. Métodos para seleccionar una muestra aleatoria.

Al seleccionar una muestra aleatoria se debe tomar en cuenta si la extracción se va a realizar con reemplazo o sin reemplazo. Como se dijo antes, en el primer caso, una vez extraído el elemento de la población este puede ser devuelto a la misma, en el segundo caso esto no es posible.
Por otro lado, dada una lista de los miembros de la población numerados del 1 al N, la extracción de los elementos que conforman la muestra se puede realizar de varias maneras entre las cuales podemos mencionar: Método del bingo, Tabla de Números aleatorios y generación de números pseudoaletorios.

1. Método del bingo. Consiste en etiquetar N papeles, bolas o cualquier otro objeto del 1 al N e introducirlas en una urna o bolsa y agitarla hasta que queden bien mezcladas, luego extraer una a la vez hasta que hayamos seleccionado n artículos donde n es el tamaño deseado de la muestra. Los miembros de la población que correspondan a los números de los artículos extraídos se incluidos en la muestra, y las características de estas unidades se mide u observan. Si la población es bastante grande, este método mecánico de selección aleatoria puede ser difícil o prácticamente imposible de implementar. Esto nos lleva a la consideración de la tabla de números aleatorios.

2. Generación de números pseudoaletorios. Existen métodos más eficaces para generar números aleatorios, en muchos de los cuales se utilizan calculadoras o computadoras. La mayoría de los paquetes estadísticos generan numeros pseudoaleatorios y en excel usando la función aleatorio() se pueden generar dichos números.

3. Tabla de Números aleatorios. Las Tablas de Números Aleatorios contienen los dígitos 0, 1, 2,..., 7, 8, 9. Tales dígitos se pueden leer individualmente o en grupos y en cualquier orden, en columnas hacia abajo, columnas hacia arriba, en fila, diagonalmente, etc., y es posible considerarlos como aleatorios. Las tablas se caracterizan por dos cosas que las hacen particularmente útiles para el muestreo al azar. Una característica es que los dígitos están ordenados de tal manera que la probabilidad de que aparezca cualquiera en un punto dado de una secuencia es igual a la probabilidad de que ocurra cualquier otro. La otra es que las combinaciones de dígitos tienen la misma probabilidad de ocurrir que las otras combinaciones de un número igual de dígitos. Estas dos condiciones satisfacen los requisitos necesarios para el muestreo aleatorio, establecidos anteriormente. La primera condición significa que en una secuencia de números, la probabilidad de que aparezca cualquier dígito en cualquier punto de la secuencia es 1/10. La segunda condición significa que todas las combinaciones de dos dígitos son igualmente probables, del mismo modo que todas las combinaciones de tres dígitos, y así sucesivamente.

   Para utilizar una Tabla de Números Aleatorios:

   a) Hacer una lista de los elementos de la población.
   b) Numerar consecutivamente los elementos de la lista, empezando con el cero (0, 00, 000, etc.).
   c) Tomar los números de una Tabla de Números Aleatorios, de manera que la cantidad de dígitos de cada uno sea igual a la del último elemento numerado de su lista. De ese modo, si el último número fue 18, 56 ó 72, se deberá tomar un dígito de dos números.
   d) Omitir cualquier dígito que no corresponda con los números de la lista o que repita cifras seleccionadas anteriormente de la tabla. Continuar hasta obtener el número de observaciones deseado.
   e) Utilizar dichos números aleatorios para identificar los elementos de la lista que se habrán de incluir en la muestra.

La tabla siguiente representa un fragmento de una tabla de números aleatorios.

Tabla de Numeros Aleatorios

6017	2438	3828	2161	6601	8762	8166	3756	6483	7405	6595	8695
3268	5788	5965	4427	9227	8468	1298	4343	1346	0861	5400	5286
0632	5878	0726	5624	7813	7905	9611	3839	6226	3452	7352	9818
0372	1222	1781	0216	5798	5805	3719	3155	6336	4710	7311	5553
3132	3375	7801	2782	1500	4249	4702	1799	9587	2788	7421	3631
3213	0670	1158	0562	6208	6641	5057	1747	7559	0548	9614	6265
6075	7161	6505	0599	1398	2947	7797	0038	4414	3904	8021	5093
2009	3799	8336	8189	8441	5748	3587	9128	2088	8840	6838	5810
8964	8261	1914	4651	9081	3202	9692	5605	7902	9525	4932	9719
7080	9448	848	8331	9069	4214	3824	2350	4986	8556	5394	1971
4098	6758	9526	6559	5435	6428	6362	7876	7746	3562	1567	7828
3328	3604	7368	9744	8842	0456	6317	0218	3826	6603	4549	2501
9976	8845	6219	2593	8337	2222	7455	1587	2778	6178	6670	4229
6420	0204	3168	5283	6869	1675	0408	7816	9054	1931	1771	3513
6523	7018	0413	5606	2869	5234	5344	5181	2457	9569	6402	9317
7475	2647	8714	6275	9693	5937	0516	1304	1156	4133	3926	1961
4928	3235	0889	1701	3778	4803	3637	6609	1152	6832	9422	8956
8355	2702	0780	5091	6964	6693	7576	9651	3543	2515	6981	4808
0084	7215	6568	4753	0215	4797	2589	2416	4746	2469	2613	7049
6319	5007	4973	3050	7658	6044	3277	2416	5823	0871	2378	0150
7335	6191	6314	2974	2783	6280	8045	6139	1575	7728	4264	4703
0164	0416	8561	4309	6759	1658	1085	6807	4425	7435	5645	4685
8751	7452	7483	5945	2360	3542	7421	9632	5936	9718	3034	7107
6070	4807	2681	1311	2724	4979	6886	2426	4486	2350	1654	4411
8094	4307	6627	6067	2654	2265	9557	4753	3174	2253	1168	2303
2778	6633	6219	4301	5528	2485	3996	5792	1741	4351	5324	4159
7672	7480	2976	3952	3061	8719	4613	2271	8921	0848	8062	1366
1449	3173	4095	2528	6684	9596	4762	1133	1784	9004	9366	1677
2984	3961	0226	3491	5758	6907	6856	1359	2532	8928	2850	3798

Para ilustrar el uso de la tabla de números aleatorios se dará el siguiente ejemplo:

Ejemplo 7.1 *suponga que tenemos 40 latas de refrescos, y que deseamos tomar una muestra de tamaño* $n = 4$ *para estudiar su condición. Nuestro primer paso es numerar las latas de 1 a 40 o apilarlas en algún orden de tal forma que puedan ser identificadas. En la tabla de números aleatorios, los dígitos deben escogerse de a dos a la vez porque la población de tamaño* $N = 40$ *es un número de dos dígitos. Luego se selecciona arbitrariamente una fila y una columna de la tabla. Suponga que la selección es fila 6, y la columna 4. Leemos los pares de dígitos a partir de la columna 4 y moviéndonos hacia la derecha, ignorando los números mayores que 40 y también cualquier número repetido cuando aparezca una segunda vez. Se continúa* *leyendo pares de dígitos hasta que cuatro unidades diferentes hayan sido seleccionadas, es decir lo numeros 05, 20, 08 y 17. Por lo tanto, las latas con la etiqueta correspondiente a dichos numeros constituyen la muestra.*

Ahora bien, si se tiene una población de tamaño N y de la ella se toma una muestra de tamaño n, obteniéndose los siguientes resultados: $x_1, \ldots, x_n$, es claro que estos resultados son algunos de los valores que pueden obtenerse cada vez que se toma una muestra de tamaño n. Por tanto, se pueden representar mediante n variables aleatorias: $X_1, \ldots, X_n$. A este conjunto de variables aleatorias se le denomina **muestra aleatoria**.

Definición 7.7 (Muestra aleatoria) *A las variables aleatorias $X_1, \ldots, X_n$ se les llama muestra aleatoria si las mismas constituyen un conjunto de n variables aleatorias independientes y provienen de la misma población, es decir, tienen la misma función de probabilidad (idénticamente distribuidas).*

Toda población está caracterizada por unos valores, denominados parámetros. estos valores poblacionales son en general, desconocidos. Para obtener una estimación de los mismos, uno de los procedimientos es la **estimación puntual**, el cual consiste en usar funciones de la muestra seleccionada, denominadas estadísticos, para aproximar al parámetro. Ese estadístico recibe el nombre de **estimador puntual**.
Sin embargo, no todo estadístico puede ser usado para estimar los parámetros de una población. Esto es, no todo estadístico es un estimador. Adicionalmente, pueden existir varios estimadores para un mismo parámetro, de los cuales se escoge el mejor.

Definición 7.8 (Estimador Puntual) *Función de la muestra aleatoria (estadístico) utilizada para aproximar el parámetro. Estadístico que cumple con ciertas propiedades que lo califican como función apropiada para estimar el parámetro.*

Definición 7.9 (Error muestral o error de muestreo) *Es el error que se comete debido al hecho dar conclusiones sobre cierta realidad, a partir de la observación de sólo una parte de ella. Es decir, es la diferencia entre el parámetro de la población y el estadístico de la muestra utilizado para estimar el parámetro.*

Definición 7.10 (Distribución muestral) *Es la distribución de probabilidad de un estadístico obtenida como resultado de la selección de un número infinito de muestras aleatorias independientes, cada una de tamaño n provenientes de la población de interés.*

7.3. Distribución muestral de la media muestral

Definición 7.11 (Media muestral) *Sea $X_1, \ldots, X_n$ una muestra aleatoria de tamaño n de una población con función de densidad $f_X(x)$ con media μ y varianza σ^2. La media muestral representada por $\widehat{\mu}$, es*

$$\widehat{\mu} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{7.1}$$

es decir, la media aritmética de los elementos de la muestra, $\widehat{\mu} = \overline{X}$.

Definición 7.12 (Esperanza matemática de la media muestral) *Sea $\overline{x}$ la media de una muestra aleatoria de tamaño n de una población con función de densidad $f_X(x)$ con media μ y varianza σ^2. Entonces*

$$E(\overline{x}) = \mu \tag{7.2}$$

Definición 7.13 (Desviación estándar de la media muestral) *Sea $\overline{X}$ la media de una muestra aleatoria de tamaño n de una población con función de densidad $f_X(x)$ con media μ y varianza σ^2. Entonces*

$$DE(\overline{x}) = \sigma_{\overline{x}} \quad = \quad \begin{cases} \frac{\sigma}{\sqrt{n}}, & \text{para poblaciones infinitas;} \\ \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}, & \text{para poblaciones finitas.} \end{cases}$$

Al factor $\sqrt{\frac{N-n}{N-1}}$ se le conoce como ***factor de corrección para poblaciones finitas*** *y debe considerarse siempre y cuando $\frac{n}{N} \geq 0{,}05$.*

Ejemplo 7.2 *Sea el conjunto de valores $\{2, 3, 5, 7, 8\}$. Si se seleccionan muestras aleatorias de tamaño 2, obtener la distribución muestral de la media muestral.*
En la tabla 7.1 se muestran las 10 posibles muestras aleatorias diferentes asociadas con estos datos. La media poblacional es $\mu = 5$ y la desviación estándar poblacional es $\sigma = 2{,}28$. Obsérvese que en la tabla 7.1 hay 8 muestras aleatorias cuya media es diferente de μ. A esa diferencia entre el valor muestral y el valor del parámetro se le denomina ***error muestral****. Por supuesto que en la práctica su valor nunca se puede calcular, pues el valor del parámetro se desconoce.*
Como se describió en el capitulo 5, la distribución de probabilidades de una variable aleatoria es el conjunto de posibles valores de la misma y sus probabilidades. De esta forma, la distribución muestral de la variable media muestral, es el conjunto de valores que toma dicha variable y sus probabilidades, respectivamente.
En la tabla 7.2 se muestra la distribución de probabilidad de la media muestral para el ejemplo 7.2

Tabla 7.1: Todas las muestras posibles para los datos del ejemplo 7.2

Muestra Nº	*Observaciones*	*Media*	
1	$2y3$	2,5	
2	$2y5$	3,5	
3	$2y7$	4,5	
4	$2y8$	5	
5	$3y5$	4	
6	$3y7$	5	
7	$3y8$	5,5	
8	$5y7$	6	
9	$5y8$	6,5	
10	$7y8$	7,5	

Tabla 7.2: Distribución muestral de $\overline{X}$ para el ejemplo 7.2

$\overline{x}_i$	$P(\overline{X} = \overline{x}_i)$	
2,5	$\frac{1}{10}$	
3,5	$\frac{1}{10}$	
4	$\frac{1}{10}$	
4,5	$\frac{1}{10}$	
5	$\frac{2}{10}$	
5,5	$\frac{1}{10}$	
6	$\frac{1}{10}$	
6,5	$\frac{1}{10}$	
7,5	$\frac{1}{10}$	

Ahora bien, construir la distribución muestral de la media muestral requiere del conocimiento de todas y cada una de las posibles muestras que se pueden extraer de la población. Esto puede resultar poco práctico o imposible de realizar. Por tanto, para la determinación de la forma de la distribución de $\overline{X}$ se deben examinar otras alternativas.

Dos casos se pueden evaluar: (1) la población tiene distribución normal y (2) la población no tiene distribución normal. En el primer caso, sin importar el tamaño de muestra, la distribución de la media muestral es normal. Como se vio en el capitulo 6; cualquier combinación lineal de variables aleatorias independientes y con distribución normal, también sigue una distribución normal. En el caso (2), si el tamaño de muestra es lo suficientemente grande, el **Teorema de Límite Central** ayuda a determinar la distribución de la media muestral.

Teorema 7.1 *Sea* $X_1, \ldots, X_n$*, una muestra aleatoria consistente de* n *variables aleatorias independientes normalmente distribuidas con media* $E(xi) = \mu$ *y varianza* $Var(xi) = \sigma^2$*,* $i =$

$1, \ldots, n$. Entonces la distribución de la media muestral $\overline{X}$, es normal con media μ y varianza $\frac{\sigma^2}{n}$

Ahora bien, si σ^2 es conocida, la variable aleatoria

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \tag{7.3}$$

sigue una distribución normal con media 0 y varianza 1. Si σ^2 es desconocida, dos situaciones pueden presentarse de acuerdo al tamaño muestral, n. Si $n \geq 30$, σ^2 se aproxima por S^2, la varianza muestral, y el resultado mostrado en la ecuación 7.6 es valido. En caso contrario, es decir, si $n < 30$, entonces

$$t = \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \tag{7.4}$$

sigue una distribución t-student con $n - 1$ grados de libertad.

Ejemplo 7.3 *Un fabricante especifica que la resistencia X de sus bolsas tiene distribución normal con media 64 litros y desviación estándar 4 litros. Calcule la probabilidad que una muestra aleatoria de 9 bolsas tenga una resistencia media menor a 60 litros.*
$X \sim N(64, 6)$, por teorema 7.1 se tiene que $\overline{X} \sim N(64, 2)$. Dado que σ^2 es conocida, se cumple que $Z = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ sigue una distribución normal estándar. Luego,

$$\begin{aligned} P(\overline{X} \leq 60) &= P\left(Z \leq \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(Z \leq \frac{60 - 64}{\frac{6}{3}}\right) \\ &= P(Z \leq -2) = 0{,}0228 \end{aligned}$$

Ejemplo 7.4 *Supóngase que en el ejemplo 7.3 la varianza poblacional es desconocida. Calcule la probabilidad que una muestra aleatoria de 30 bolsas tenga una resistencia media menor a 63 litros, si la desviación estándar muestral es $s = 6{,}5$.*
Dado que $n \geq 30$, la probabilidad solicitada se puede calcular como en el ejemplo 7.3.

$$\begin{aligned} P(\overline{X} \leq 63) &= P\left(Z \leq \frac{\overline{X} - \mu}{\frac{s}{\sqrt{n}}}\right) = P\left(Z \leq \frac{63 - 64}{\frac{6{,}5}{\sqrt{30}}}\right) \\ &= P(Z \leq -0{,}84) = 0{,}2005 \end{aligned}$$

Ejemplo 7.5 *En el ejemplo 7.4, calcular la probabilidad que una muestra aleatoria de 25 bolsas tenga una resistencia media menor a 66 litros.*

Dado que $n = 25 < 30$, la probabilidad solicitada se puede calcular usando la distribución t-student.

$$\begin{aligned} P(\overline{X} \leq 62) &= P\left(T_{n-1} \leq \frac{\overline{X} - \mu}{\frac{s}{\sqrt{n}}}\right) = P\left(T_{19} \leq \frac{66 - 64}{\frac{6{,}5}{\sqrt{20}}}\right) \\ &= P(T_{19} \leq 1{,}376) = 0{,}0946 \end{aligned}$$

Teorema 7.2 (Teorema de Límite Central (TLC)) *Sean $X_1, \ldots, X_n$, n variables aleatorias independientes idénticamente distribuidas con media μ y varianza σ^2,ambas finitas. Si n es lo suficientemente grande, entonces*

$$\overline{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$$

En general, se acepta como una aproximación apropiada al modelo normal siempre que $n \geq 30$ y se dice que la muestra es "suficientemente grande". En este caso, si se desconoce la varianza de la población, la misma se puede aproximar mediante la varianza muestral, S^2.

Ejemplo 7.6 *Estudios anteriores muestran que el ingreso promedio de las familias en la ciudad de mérida es de 1500 BsF. Cual es la probabilidad de que el ingreso promedio de una muestra aleatoria de 100 familias de dicha ciudad se encuentre entre 1450 y 1600 BsF, si la desviación estándar muestral es igual a 500 BsF.*
En este caso, X es la variable aleatoria ïngreso de las familias en la ciudad de Mérida". Es una variable aleatoria cuya distribución de probabilidad es desconocida, con media igual a 1500 y varianza desconocida. Por el Teorema 7.2, $\overline{X}$ tiene distribución aproximadamente normal cuando n es lo suficientemente grande, en general, $n \geq 30$. Por tanto, $\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ tiene distribución aproximadamente normal estándar, aproximando a σ por S. De esta forma,

$$\begin{aligned} P(1450 \leq \overline{X} \leq 1600) &= P(\overline{X} \leq 1600) - P(\overline{X} \leq 1450) \\ &= P\left(Z \leq \frac{1600 - 1500}{\frac{500}{\sqrt{100}}}\right) - P\left(Z \leq \frac{1450 - 1500}{\frac{500}{\sqrt{100}}}\right) \\ &= P(Z \leq 2) - P(Z \leq -1) = 0{,}9772 - 0{,}1587 = 0{,}8185 \end{aligned}$$

Aun, si las variables no son idénticamente distribuidas, se podría demostrar que $\overline{X}$ sigue una distribución normal con media $\frac{\sum \mu_i}{n}$ y varianza $\frac{\sum \sigma_i^2}{n^2}$

7.4. Distribución muestral de la diferencia de medias muestrales

Teorema 7.3 *Sea $X_1, \ldots, X_{n_1}$, una muestra aleatoria de tamaño n_1 procedente de una población normal con media μ_1 y varianza σ_1^2, $i = 1, \ldots, n_1$. Sea $X_1, \ldots, X_{n_2}$, una muestra aleatoria de* *tamaño n_2 procedente de una población normal con media μ_2 y varianza σ_2^2, $i = 1, \ldots, n_2$. Si las muestras son independientes, entonces $\Delta\overline{X} = (\overline{X}_1 - \overline{X}_2)$ sigue una distribución normal con media $\Delta\mu = (\mu_1 - \mu_2)$ y varianza $\sigma^2_{\Delta\overline{X}} = \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$.*

Si σ_1^2 y σ_2^2 son conocidas, la variable aleatoria

$$Z = \frac{\Delta\overline{X} - \Delta\mu}{\sigma_{\Delta\overline{X}}} \tag{7.5}$$

sigue una distribución normal con media 0 y varianza 1. Si σ_1^2 y σ_2^2 son desconocidas, dos situaciones pueden presentarse de acuerdo a los tamaños muestrales, n_1 y n_2. Si $n_1 \geq 30$ y $n_2 \geq 30$, σ_1^2 y σ_2^2 se aproximan por S_1^2 y S_2^2, las varianzas muestrales, y el resultado mostrado en la ecuación 7.5 es valido. En caso contrario, es decir, si al menos un $n_i < 30$, entonces

$$t = \frac{\Delta\overline{X} - \Delta\mu}{S_{\Delta\overline{X}}} \tag{7.6}$$

sigue una distribución t-student con δ grados de libertad. La expresión de $S_{\Delta\overline{X}}$ y de los grados de libertad dependen de la suposición que se haga sobre σ_1^2 y σ_2^2.

$$\sigma_1^2 = \sigma_2^2 \quad \Rightarrow \quad \begin{cases} \delta = n_1 + n_2 - 2, \\ S^2_{\Delta\overline{X}} = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} \end{cases} \tag{7.7}$$

$$\sigma_1^2 \neq \sigma_2^2 \quad \Rightarrow \quad \begin{cases} \delta = \dfrac{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{(n_1-1)} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{(n_2-1)}}, \\ S^2_{\Delta\overline{X}} = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} \end{cases} \tag{7.8}$$

Existen situaciones en las que las poblaciones muestreadas no siguen una distribución normal. Una extensión del teorema 7.2 permite, dar solución en este caso, siempre y cuando tanto n_1 como n_2, sean lo suficientemente grandes.

Ejemplo 7.7 *La edad X_1 de los empleados hombres de una empresa sigue una distribución normal con media 28 y varianza 81, mientras que la edad X_2 de las empleadas sigue una distribución normal con media 24 y varianza 81. Si se seleccionan muestras de tamaño 20 para cada grupo de empleados, calcular a probabilidad de que la diferencia entre las dos medias muestrales sea mayor o igual a 5 años.*

Obsérvese que $X_1 \sim N(28; 9)$ y $X_2 \sim N(24; 9)$. Por tanto, usando el teorema 7.3 se tiene que

$\Delta\overline{X} \sim N(4; 2,85)$. *Luego*

$$\begin{aligned} P(\Delta\overline{X} \geq 5) &= 1 - P\left(Z \leq \frac{5-4}{2{,}85}\right) \\ &= 1 - P(Z \leq 0{,}35) = 1 - 0{,}6368 = 0{,}3632 \end{aligned}$$

Ejemplo 7.8 *Supóngase que en el ejemplo 7.7 las distribuciones de las poblaciones así como sus varianzas son desconocidas. Si seleccionan ahora muestras de tamaño* $n_1 = 36$ *y* $n_2 = 49$*, calcular a probabilidad de que la diferencia entre las dos medias muestrales sea mayor o igual a 5 años, si las varianzas muestrales son* $S_1^2 = 81$ *y* $S_2^2 = 100$.
Dado que $n_1 \geq 30$ *y* $n_2 \geq 30$*, de acuerdo al teorema 7.2 se tiene que* $\Delta\overline{X} \sim N(4; 2,85)$. *Luego*

$$\begin{aligned} P(\Delta\overline{X} \geq 5) &= 1 - P\left(Z \leq \frac{5-4}{2{,}07}\right) \\ &= 1 - P(Z \leq 0{,}48) = 1 - 0{,}6844 = 0{,}3156 \end{aligned}$$

Ejemplo 7.9 *Resolver el ejemplo 7.7 suponiendo que las varianzas poblacionales son desconocidas pero iguales y que sus estimaciones están dadas por* $S_1^2 = 16$ *y* $S_2^2 = 16$.
Como las poblaciones son normales, las varianzas poblacionales desconocidas y los tamaños muestrales menores a 30. *Entonces, usando la ecuación 7.7 se tiene que,*

$$\begin{aligned} P(\Delta\overline{X} \geq 5) &= 1 - P\left(t_{38} \leq \frac{5-4}{3{,}54}\right) \\ &= 1 - P(t_{38} \leq 0{,}25) = 1 - 0{,}4 = 0{,}6 \end{aligned}$$

Ejemplo 7.10 *Resolver el ejemplo 7.7 suponiendo que las varianzas poblacionales son desconocidas pero diferentes y que sus estimaciones están dadas por* $S_1^2 = 16$ *y* $S_2^2 = 16$.
Como las poblaciones son normales, las varianzas poblacionales desconocidas y los tamaños muestrales menores a 30. *Entonces, usando la ecuación 7.8 se tiene que,*

$$\begin{aligned} P(\Delta\overline{X} \geq 5) &= 1 - P\left(t_{24} \leq \frac{5-4}{1{,}27}\right) \\ &= 1 - P(t_{24} \leq 0{,}79) = 1 - 0{,}2185 = 0{,}7815 \end{aligned}$$

7.5. Distribución muestral de varianzas muestrales

Teorema 7.4 *Sea* $X_1, \ldots, X_n$*, una muestra aleatoria consistente de* n *variables aleatorias independientes procedentes de una población normal con media* μ *y varianza* σ^2*,* $i = 1, \ldots, n$.

Entonces

$$\frac{\sum_{i=1}^{n}(x_i - \overline{X})^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \tag{7.9}$$

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

Ejemplo 7.11 *Una muestra aleatoria de tamaño* 10 *es tomada de una población distribuida normalmente con media* 8 *y varianza* 9*. Calcular la probabilidad de que la varianza muestral sea mayor que* 3,753
El calculo de esta probabilidad se hace usando el teorema 7.4.

$$\begin{aligned} P(S^2 \leq 3{,}753) &= P\left(\chi_9^2 \leq \frac{(10-1)3{,}753}{9}\right) \\ &= P(\chi_9^2 \leq 3{,}753) = 0{,}9269 \end{aligned}$$

Teorema 7.5 *Sea* $X_1, \ldots, X_n$*, una muestra aleatoria consistente de* n *variables aleatorias independientes procedentes de una población normal con media* μ_1 *y varianza* σ_1^2, $i = 1, \ldots, n$*. Sea* $Y_1, \ldots, Y_m$*, una muestra aleatoria consistente de* m *variables aleatorias independientes procedentes de una población normal con media* μ_2 *y varianza* σ_2^2, $i = 1, \ldots, m$*. Si las muestras son independientes, entonces el cociente*

$$\frac{\frac{(n-1)S_1^2}{\sigma_1^2}}{\frac{(m-1)S_2^2}{\sigma_2^2}} \tag{7.10}$$

sigue una distribución F *con* $(n-1)$ *y* $(m-1)$ *grados de libertad.*

Este cociente es de especial interés cuando se desea hacer inferencia acerca de las magnitudes relativas de σ_1^2 y σ_2^2.

Ejemplo 7.12 *Sean* $X_1 \sim N(\mu_1; 10)$ *y* $X_2 \sim N(\mu_2; 15)$*, variables aleatorias independientes. Se toman muestras de tamaño* $n = 16$ *y* $m = 16$*, respectivamente. Calcular la probabilidad de que* S_1^2 *sea a lo sumo igual a* S_2^2*.*

$$\begin{aligned} P\left(\frac{S_1^2}{S_2^2} \leq 1\right) &= P\left(\frac{\frac{(n-1)S_1^2}{\sigma_1^2}}{\frac{(m-1)S_2^2}{\sigma_2^2}} \leq \frac{(n-1)\sigma_2^2}{(m-1)\sigma_1^2}\right) \\ &= P(F_{15;15} \leq 2{,}25) = 0{,}067 \end{aligned}$$

Dado que el valor 2,25 *no está tabulado, debe obtenerse por interpolación, es decir*

$$\begin{aligned} y &= y_a + (x - x_a)\frac{(y_b - y_a)}{(x_b - x_a)} \\ &= 0{,}10 + (2{,}25 - 1{,}97)\frac{(0{,}05 - 0{,}10)}{(2{,}40 - 1{,}97)} \\ &= 0{,}06744 \end{aligned}$$

7.6. Distribución muestral de la proporción muestral

Definición 7.14 (Proporción muestral) *La proporción muestral se define como el cociente del número de elementos de la muestra que tienen la característica deseada, entre el número total de elementos de la muestra. Esto es, si X representa el número de elementos de la muestra que tienen la característica deseada y n es el número total de elementos de la muestra, entonces*

$$p = \frac{X}{n} \tag{7.11}$$

La esperanza matemática y la varianza de $\widehat{p}$ están dadas por:

$$E(p) = E\left(\frac{X}{n}\right) = \pi;$$

y

$$\sigma_p^2 = \begin{cases} \frac{\pi(1-\pi)}{n}, & \text{para poblaciones infinitas;} \\ \frac{\pi(1-\pi)}{n}\frac{N-n}{N-1}, & \text{para poblaciones finitas.} \end{cases}$$

Ahora bien, X es una variable aleatoria binomial. Al ser n constante, la probabilidad de $\frac{x}{n}$ es la misma de x. De esta forma, la distribución muestral de p es una distribución de probabilidad discreta. En el capitulo 6 se mostró como una distribución binomial se aproxima mediante una distribución normal, siempre y cuando n sea lo suficientemente grande y se satisfaga que $np \geq 5$ y $n(1-p) \geq 5$.

Teorema 7.6 *Sea $X_1, \ldots, X_n$, una muestra aleatoria consistente de n variables aleatorias independientes de una población binomial con parámetro π. Sea p la proporción de elementos en la muestra que poseen una característica de interés. Si se cumple que $np \geq 5$ y $n(1-p) \geq 5$, entonces la distribución de p, es normal con media π y varianza $\frac{\pi(1-\pi)}{n}$.*

Si la población es finita entonces a la expresión de la varianza se le debe agregar el factor de corrección para poblaciones finitas, $\frac{N-n}{N-1}$.

Ejemplo 7.13 *Se toma una muestra aleatoria de tamaño* 16 *de una población binomial con* $\pi = 0{,}7$*. Suponiendo que la población es infinita, calcular la probabilidad de que la proporción muestral sea menor que* 0,6*.*
Dado que $np \geq 5$ *y* $n(1-p) \geq 5$*, se puede usar el resultado mostrado en el teorema 7.6. De esta forma,* $p \sim N(0{,}7; 0{,}0525)$*.*

$$\begin{aligned} P(p < 0{,}6) &= P\left(Z < \frac{0{,}6 - 0{,}7)}{0{,}0525}\right) \\ &= P(Z < -1,904761905) = 0{,}0281 \end{aligned}$$

7.7. Distribución muestral de la diferencia de proporciones muestrales

Muchas aplicaciones involucran poblaciones de datos cualitativos que deben compararse utilizando proporciones o porcentajes. Determinar si la proporción de los estudiantes que aprueban matemáticas es mayor que las de los que aprueban inglés o, verificar si existe diferencia entre la proporción de artículos defectuosos que genera la máquina A a los que genera la máquina B, son ejemplos de estas aplicaciones.

Teorema 7.7 *Sean* $X_1, \ldots, X_n$ *y* $Y_1, \ldots, Y_m$*, muestra aleatorias independientes procedentes de dos poblaciones binomiales. Entonces, para tamaños de muestra grande tales que se satisfacen* $n_1p_1 \geq 5$*,* $n_1(1-p_1) \geq 5$*,* $n_2p_2 \geq 5$ *y* $n_2(1-p_2) \geq 5$*, la distribución muestral de diferencia de proporciones* $\Delta p = (p_1 - p_2)$*, es aproximadamente normal con media* $\Delta\pi = (\pi_1 - \pi_2)$ *y varianza* $\left(\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}\right)$*.*

Ejemplo 7.14 *Los hombres y mujeres adultos de una ciudad grande de Venezuela difieren en sus opiniones sobre el el gobierno. Se cree que el* 39 % *de los hombres adultos califican el gobierno como bueno, mientras que el* 37 % *de las mujeres adultas así lo creen. Si se pregunta a dos muestras aleatorias de* 50 *hombres y* 50 *mujeres su opinión sobre el gobierno, determinar la probabilidad de que el porcentaje de hombres a favor sea al menos* 4 % *mayor que el de las mujeres.*

$$\begin{aligned} P(\Delta p \geq 0{,}04) &= 1 - P\left(Z < \frac{0{,}04 - 0{,}02)}{0,00942}\right) \\ &= 1 - P(Z < 2{,}12) = 1 - 0{,}9830 = 0{,}017 \end{aligned}$$

Ejemplo 7.15 *En el ejemplo 7.14, calcular la probabilidad de que dicha diferencia sea a lo sumo* *sea del 3 %.*

$$\begin{aligned} P(\Delta p \leq 0{,}03) &= P\left(Z < \frac{0{,}03 - 0{,}02)}{0,00942}\right) \\ &= P(Z < 1{,}06) = 0{,}8554 \end{aligned}$$

7.8. Ejercicios

1. Considérese la variable aleatoria X cuya función de probabilidad está dada por: Obtener la

x	0	1	2	3
$P(X = x)$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{3}{10}$	$\frac{2}{10}$

   distribución de probabilidad de:

   a) $\overline{X}$

   b) $X_1 + X_2$

   c) S^2

2. En una población se presenta una característica de interés en una proporción $\pi = 0{,}3$ de los individuos que la componen. Definimos una variable aleatoria X que toma el valor 1 para aquellos individuos que poseen la característica y 0 para los que no.

   a) Escriba la distribución poblacional de X.

   b) Si p es la proporción de veces que aparece el valor 1 en muestras aleatorias simples de tamaño 3, obtener la distribución en el muestreo de p.

3. Sea X una variable aleatoria con distribución $N(90; 20)$. Si se toma una muestra aleatoria de tamaño 16. Calcular la probabilidad de que la media muestral:

   a) sea mayor o igual que 92.

   b) sea menor o igual que 98.

   c) se encuentre entre 85 y 100.

4. La pesca diaria en kilogramos de un pescador de merluza sigue una distribución normal con una media de 12 kilos y una desviación estándar de 2 kilos. Si se toma una muestra de 25 trampas para merluzas:

   a) cual es la probabilidad de que la pesca promedio diaria por trampa sea como máximo 12 kilos.

   b) Cual es la probabilidad de que la pesca promedio diaria por trampa este entre 13 y 16 kilos.

5. El Coeficiente intelectual de los adultos de cierta población sigue una distribución normal con media igual a 100 y desviación estándar igual 16. Se toma una muestra aleatoria de 25 individuos. Calcular la probabilidad de que la media muestral:

   a) sea inferior a 95.

   b) sea superior a 95.

c) este entre 96 y 104

6. El tiempo que tardan los habitantes de Ejido para llegar a la ciudad de Mérida es una variable aleatoria que sigue una distribución normal con media igual 57 minutos. Se toma una muestra aleatoria de de 16 habitantes. Calcular la probabilidad de que la media muestral sea menor a 60 minutos, si la desviación estándar muestral es igual a 15 minutos.

7. El incremento porcentual de los sueldos de los funcionarios públicos tiene una media de 25 % y una desviación típica de 4 %. Si se toma una muestra aleatoria de 100 empleados, calcular la probabilidad de que el incremento medio muestral porcentual sea menor del 20 %.

8. Se toma una muestra aleatoria de 50 casas nuevas en la ciudad de Mérida y se les registra su valor de venta. Calcular la probabilidad de que el valor promedio de venta de las 50 casas se encuentre entre 600000 y 1000000 bolívares fuertes si la desviación estándar muestral fue de 60000 bolívares fuertes.

9. Sean $X_1 \sim N(50; 2)$ y $X_2 \sim N(75; 3)$, variables aleatorias independientes. Si se toman muestras de tamaño $n_1 = 20$ y $n_2 = 20$, calcular la probabilidad de que $\overline{X}_1$ supere a $\overline{X}_2$ en 5 unidades o menos.

10. En el ejercicio 9, calcular:

    a) $P(\Delta\overline{X} \geq 7)$.
    b) $P(10 \leq \Delta\overline{X} \leq 15)$.
    c) $P(5 \leq \Delta\overline{X} \leq 12)$.

11. Sean $X_1 \sim N(1000; \sigma_1)$ y $X_2 \sim N(1500; \sigma_2)$, variables aleatorias independientes. Se toman muestras de tamaño $n_1 = 50$ y $n_2 = 50$, respectivamente. Si $s_1^2 = 50$ y $s_1^2 = 70$ calcular:

    a) $P(\Delta\overline{X} \leq 500)$.
    b) $P(\Delta\overline{X} \geq 500)$.
    c) $P(400 \leq \Delta\overline{X} \leq 600)$.

12. Cierta institución desea comparar el rendimiento estudiantil de las hembras, X_1, y el rendimiento estudiantil de los varones, X_2. Para ello se toman muestras aleatoria de tamaño $n_1 = 30$ y $n_2 = 30$, respectivamente, las cuales producen $s_1^2 = 2$ y $s_1^2 = 4$. Si se sabe que $\mu_{X_1} = 15$ y $\mu_{X_1} = 14$, calcular:

    a) $P(\Delta\overline{X} \leq 0)$.
    b) $P(\Delta\overline{X} \geq 2)$.
    c) $P(2 \leq \Delta\overline{X} \leq 4)$.

13. En el ejercicio 12, supóngase que $X_1 \sim N(15; \sigma_1)$, $X_2 \sim N(14; \sigma_2)$, $n_1 = 20$ y $n_2 = 20$. Suponiendo además que $\sigma_1^2 = \sigma_1^2$, calcular:

   a) $P(\Delta\overline{X} \leq 2)$.
   b) $P(\Delta\overline{X} \geq 1)$.
   c) $P(1 \leq \Delta\overline{X} \leq 3)$.

14. Resolver el ejercicio 13, suponiendo que $\sigma_1^2 \neq \sigma_1^2$.

15. Se ha estimado que el 25 % de los egresados de FACES consideran que la ética es fundamental en el buen desempeño de sus labores. Se toma una muestra de 80 egresados. Calcular la probabilidad de que más de la mitad de ellos opinen de esta manera.

16. Se toma una muestra de 100 edificaciones en la ciudad de Mérida para estimar la proporción de las mismas en dicha ciudad. Supongamos que el 30 % de todas las edificaciones son antiguas. Hallar la probabilidad de que la proporción de edificaciones antiguas en la ciudad de Mérida esté entre 0,25 y 0,35.

17. Se desea averiguar cuántos profesores universitarios en Venezuela están a favor de un cambio en el sistema educativo de dicho país. Se selecciona en forma aleatoria un grupo de 80 de estos profesores. Si el 55 % de todos los profesores de la población opina que un cambio resultaría beneficioso, ¿cuál es la probabilidad de que: la mitad de los individuos de la muestra tengan es opinión?

   a) la mitad de los individuos de la muestra tengan es opinión?.
   b) menos de la mitad de los individuos de la muestra tengan es opinión?.
   c) más de la mitad de los individuos de la muestra tengan es opinión?.

18. La proporción de estudiantes varones con vehículo es $\pi_1 = 0{,}18$, mientras que la de mujeres es $\pi_2 = 0{,}20$. Se toman muestras aleatorias de tamaño $n_1 = 30$ y $n_2 = 40$, para varones y hembras respectivamente. Calcular

   a) $P(\Delta p \leq 0{,}1)$
   b) $P(\Delta p > 0{,}1)$
   c) $P(0{,}1 \leq \Delta p \leq 0{,}3)$

19. La proporción de obreros y empleados administrativos que faltan a su trabajo es de $\pi_1 = 0{,}25$ y $\pi_2 = 0{,}35$, respectivamente. Si se toman muestras de tamaño 150 para grupo de trabajadores, calcular:

   a) $P(\Delta p \leq -0{,}1)$
   b) $P(\Delta p > -0{,}05)$
   c) $P(-0{,}01 \leq \Delta p \leq -0{,}15)$

20. El tiempo que requiere cierto equipo de fútbol para marcar su primer gol es una variable aleatoria con distribución normal con una desviación estándar de 10 minutos. Si se observan aleatoriamente 17 partidos de dicho equipo, encuentre la probabilidad de que la varianza muestral sea mayor que 15 minutos.

21. Encuentre la probabilidad de que una muestra aleatoria de 25 observaciones de una población normal con varianza igual a 9, tenga una varianza entre 7 y 12.

22. La gerencia de control de calidad de Aguas de Mérida realiza un estudio sobre la cantidad de cloro en el agua potable. Para ello se analizan diez muestras de agua seleccionadas en forma aleatoria. ¿Cual es la probabilidad de que la variabilidad en la cantidad de agua potable en la muestra sea inferior a 1,44 partes por millón?. Supóngase que la distribución de la variable cantidad de cloro en el agua potable sigue una distribución normal con varianza igual a 2,2

23. Sean $X_1 \sim N(\mu_1; 5)$ y $X_2 \sim N(\mu_2; 9)$, variables aleatorias independientes. Se toman muestras de tamaño $n_1 = 25$ y $n_2 = 25$, respectivamente. Calcular:

    a) $P\left(\frac{S_1^2}{S_2^2} \leq 0{,}8\right)$.
    b) $P\left(\frac{S_1^2}{S_2^2} \geq 0{,}6\right)$.
    c) $P\left(0{,}5 \leq \frac{S_1^2}{S_2^2} \leq 1\right)$.

24. El tiempo de trabajo de dos grupos de trabajadores son variables aleatorias independientes cuyas distribuciones son normales con varianzas $\sigma_1^2 = 9$ horas y $\sigma_2^2 = 4$ horas. Si se toman muestras aleatorias de tamaño 20 para los 2 grupos, calcular:

    a) $P(S_1^2 < S_2^2)$.
    b) $P(S_1^2 > S_2^2)$.
    c) $P(S_1^2 > 2S_2^2)$.