Diferencia de medias

Ejemplo:

En un estudio para comparar los pesos promedio de niños y niñas de sexto grado en una escuela primaria se usará una muestra aleatoria de 20 niños y otra de 25 niñas. Se sabe que tanto para niños como para niñas los pesos siguen una distribución normal. El promedio de los pesos de todos los niños de sexto grado de esa escuela es de 100 libras y su desviación estándar es de 14.142, mientras que el promedio de los pesos de todas las niñas del sexto grado de esa escuela es de 85 libras y su desviación estándar es de 12.247 libras. Si $\bar{x}_1$ representa el promedio de los pesos de 20 niños y $\bar{x}_2$ es el promedio de los pesos de una muestra de 25 niñas, encuentre la probabilidad de que el promedio de los pesos de los 20 niños sea al menos 20 libras más grande que el de las 25 niñas.

Solución:

Datos:

μ_1 = 100 libras

μ_2 = 85 libras

σ_1 = 14.142 libras

σ_2 = 12.247 libras

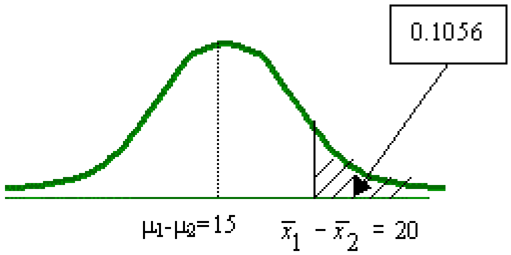
n_1 = 20 niños

n_2 = 25 niñas

$p(\bar{x}_1 - \bar{x}_2 \rangle 20)$ = ?

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{20 - (100 - 85)}{\sqrt{\frac{(14.142)^2}{20} + \frac{(12.247)^2}{25}}} = 1.25$$

Por lo tanto, la probabilidad de que el promedio de los pesos de la muestra de niños sea al menos 20 libras más grande que el de la muestra de las niñas es 0.1056.



Ejemplo:

Uno de los principales fabricantes de televisores compra los tubos de rayos catódicos a dos compañías. Los tubos de la compañía A tienen una vida media de 7.2 años con una desviación estándar de 0.8 años, mientras que los de la B tienen una vida media de 6.7 años con una desviación estándar de 0.7. Determine la probabilidad de que una muestra aleatoria de 34 tubos de la compañía A tenga una vida promedio de al menos un año más que la de una muestra aleatoria de 40 tubos de la compañía B.

Solución:

Datos:

μ_A = 7.2 años

μ_B = 6.7 años

σ_A = 0.8 años

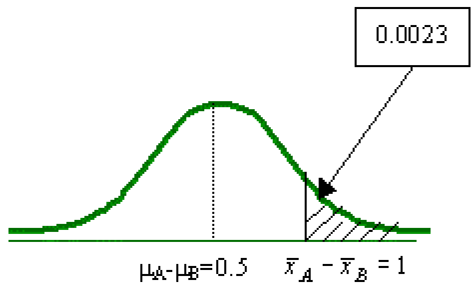
σ_B = 0.7 años

n_A = 34 tubos

n_B = 40 tubos

$p(\bar{x}_A - \bar{x}_B > 1)$ = ?

$$z = \frac{(\bar{x}_A - \bar{x}_B) - (\mu_A - \mu_B)}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} = \frac{1 - (7.2 - 6.7)}{\sqrt{\frac{(0.8)^2}{34} + \frac{(0.7)^2}{40}}} = 2.84$$



Ejemplo:

Se prueba el rendimiento en km/L de 2 tipos de gasolina, encontrándose una desviación estándar de 1.23km/L para la primera gasolina y una desviación estándar de 1.37km/L para la segunda gasolina; se prueba la primera gasolina en 35 autos y la segunda en 42 autos.

a. ¿Cuál es la probabilidad de que la primera gasolina de un rendimiento promedio mayor de 0.45km/L que la segunda gasolina?
b. ¿Cuál es la probabilidad de que la diferencia en rendimientos promedio se encuentre entre 0.65 y 0.83km/L a favor de la gasolina 1?.

Solución:

En este ejercicio no se cuenta con los parámetros de las medias en ninguna de las dos poblaciones, por lo que se supondrán que son iguales.

Datos:

σ_1 = 1.23 Km/Lto

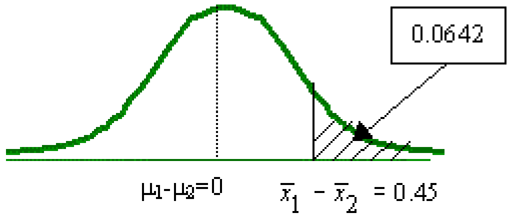
σ_2 = 1.37 Km/Lto

n_1 = 35 autos

n_2 = 42 autos

a. $p(\bar{x}_1 - \bar{x}_2 \rangle 0.45)$ = ?

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{0.45 - 0}{\sqrt{\frac{(1.23)^2}{35} + \frac{(1.37)^2}{42}}} = 1.52$$

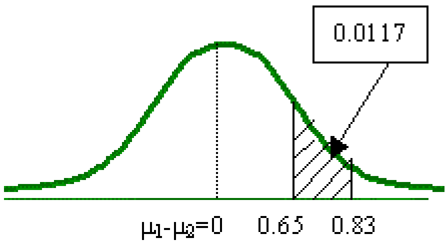


b. $p(0.65 \leq \bar{x}_1 - \bar{x}_2 \leq 0.83) =$

?

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{0.65 - 0}{\sqrt{\frac{(1.23)^2}{35} + \frac{(1.37)^2}{42}}} = 2.19$$

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{0.83 - 0}{\sqrt{\frac{(1.23)^2}{35} + \frac{(1.37)^2}{42}}} = 2.80$$



La probabilidad de que la diferencia en rendimientos promedio en las muestras se encuentre entre 0.65 y 0.83 Km/Lto a favor de la gasolina 1 es de 0.0117.

Distribución Muestral de Diferencia de Proporciones

Ejemplo:

Los hombres y mujeres adultos radicados en una ciudad grande del norte difieren en sus opiniones sobre la promulgación de la pena de muerte para personas culpables de asesinato. Se cree que el 12% de los hombres adultos están a favor de la pena de muerte, mientras que sólo 10% de las mujeres adultas lo están. Si se pregunta a dos muestras aleatorias de 100 hombres y 100 mujeres su opinión sobre la promulgación de la pena de muerte, determine la probabilidad de que el porcentaje de hombres a favor sea al menos 3% mayor que el de las mujeres.

Solución:

Datos:

$P_H = 0.12$

$P_M = 0.10$

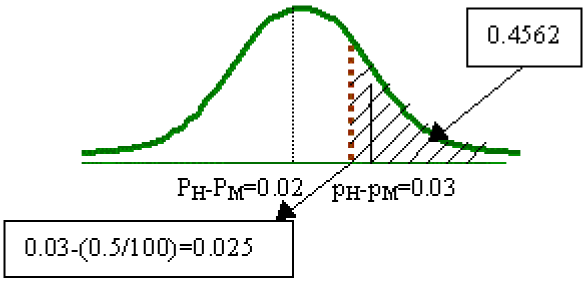
$n_H = 100$

$n_M = 100$

$p(p_H - p_M \geq 0.03) = ?$

Se recuerda que se está incluyendo el factor de corrección de 0.5 por ser una distribución binomial y se está utilizando la distribución normal.

$$z = \frac{(p_H - p_M) - (P_H - P_M)}{\sqrt{\frac{P_H q_H}{n_H} + \frac{P_M q_M}{n_M}}} = \frac{0.025 - (0.12 - 0.10)}{\sqrt{\frac{(0.12)(0.88)}{100} + \frac{(0.10)(0.90)}{100}}} = 0.11$$



Se concluye que la probabilidad de que el porcentaje de hombres a favor de la pena de muerte, al menos 3% mayor que el de mujeres es de 0.4562.

Ejemplo:

Una encuesta del Boston College constó de 320 trabajadores de Michigan que fueron despedidos entre 1979 y 1984, encontró que 20% habían estado sin trabajo durante por lo menos dos años. Supóngase que tuviera que seleccionar otra muestra aleatoria de 320 trabajadores de entre todos los empleados despedidos entre 1979 y 1984. ¿Cuál sería la probabilidad de que su porcentaje muestral de trabajadores sin empleo durante por lo menos dos años, difiera del porcentaje obtenido en la encuesta de Boston College, en 5% o más?

Solución:

En este ejercicio se cuenta únicamente con una población, de la cual se están extrayendo dos muestras y se quiere saber la probabilidad de la diferencia de los porcentajes en esas dos muestras, por lo que se debe de utilizar la distribución muestral de proporciones con $P_1= P_2$, ya que es una misma población.

Otra de las situaciones con la cual nos topamos es que desconocemos la proporción de trabajadores despedidos entre 1979 y 1984 que estuvieron desempleados por un período de por lo menos dos años, sólo se conoce la
$p_1= 0.20$ ya que al tomar una muestra de 320 trabajadores se observó esa proporción.

En la fórmula de la distribución muestral de proporciones para el cálculo de probabilidad se necesita saber las proporciones de las poblaciones, las cuales en este ejercicio las desconocemos, por lo que se utilizará el valor de 0.20 como una estimación puntual de P. En el siguiente tema se abordará el tema de estimación estadística y se comprenderá el porque estamos utilizando de esa manera el dato.

También debe de comprenderse la pregunta que nos hace este problema, ¿cuál sería la probabilidad de que su porcentaje muestral de trabajadores sin empleo durante por lo menos dos años, ***difiera*** del porcentaje obtenido en la encuesta de Boston College, en 5% o más?, la palabra ***difiera*** quiere decir que puede existir una diferencia a favor de la muestra uno, o a favor de la muestra dos, por lo que se tendrán que calcular dos áreas en la distribución y al final sumarlas.

Datos:

$p_1 = 0.20$

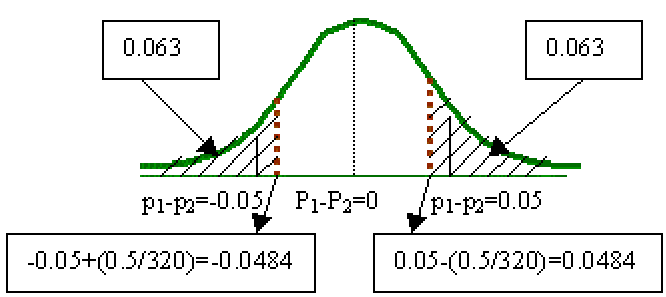
$n_1 = 320$ trabajadores

$n_2 = 320$ trabajadores

$P_1 = P_2$

$$z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{\dfrac{P_1 q_1}{n_1} + \dfrac{P_2 q_2}{n_2}}} = \frac{0.0484 - 0}{\sqrt{\dfrac{(0.20)(0.80)}{320} + \dfrac{(0.20)(0.80)}{320}}} = 1.53$$

$$z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{\dfrac{P_1 q_1}{n_1} + \dfrac{P_2 q_2}{n_2}}} = \frac{-0.0484 - 0}{\sqrt{\dfrac{(0.20)(0.80)}{320} + \dfrac{(0.20)(0.80)}{320}}} = -1.53$$



$$p(-0.05 \geq p_1 - p_2 \geq 0.05) = 0.063 + 0.063 = 0.1260$$

La probabilidad de que su proporcion muestral de trabajadores sin empleo durante por lo menos dos años, ***difiera*** del porcentaje obtenido en la encuesta de Boston College, en 0.05 o más es de 0.1260.

Ejemplo:

Se sabe que 3 de cada 6 productos fabricados por la máquina 1 son defectuosos y que 2 de cada 5 objetos fabricados por la máquina 2 son defectuosos; se toman muestras de 120 objetos de cada máquina:

a. ¿cuál es la probabilidad de que la proporción de artículos defectuosos de la máquina 2 rebase a la máquina 1 en por lo menos 0.10?
b. ¿cuál es la probabilidad de que la proporción de artículos defectuosos de la máquina 1 rebase a la máquina 2 en por lo menos 0.15?

Solución:

Datos:

$P_1 = 3/6 = 0.5$

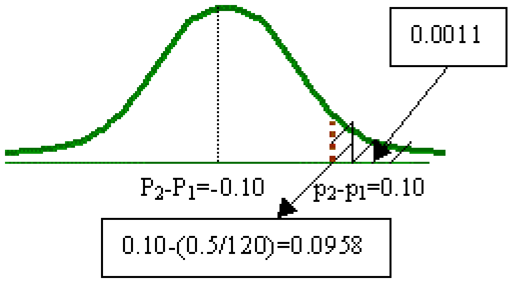
$P_2 = 2/5 = 0.4$

$n_1 = 120$ objetos

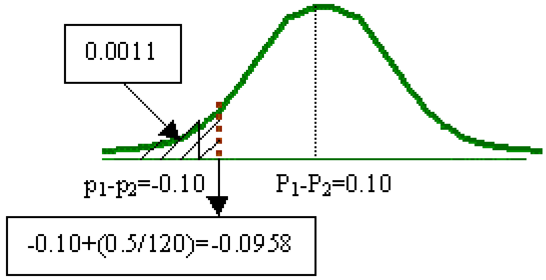
$n_2 = 120$ objetos

a. $p(p_2 - p_1 \geq 0.10) = ?$

$$z = \frac{(p_2 - p_1) - (P_2 - P_1)}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}} = \frac{0.0958 - (-0.10)}{\sqrt{\frac{(0.50)(0.50)}{120} + \frac{(0.40)(0.60)}{120}}} = 3.06$$



Otra manera de hacer este ejercicio es poner P_1-P_2:



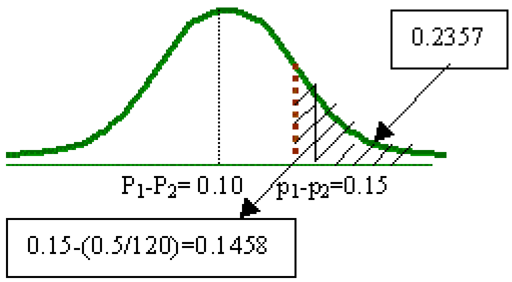
$$z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}} = \frac{-0.0958 - 0.10}{\sqrt{\frac{(0.50)(0.50)}{120} + \frac{(0.40)(0.60)}{120}}} = -3.06$$

La probabilidad de que exista una diferencia de proporciones de artículos defectuosos de por lo menos 10% a favor de la máquina 2 es de 0.0011.

b. $p(p_1\text{-}p_2 \geq$

0.15)=?

$$z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}} = \frac{0.1458 - 0.10}{\sqrt{\frac{(0.50)(0.50)}{120} + \frac{(0.40)(0.60)}{120}}} = 0.72$$



La probabilidad de que exista una diferencia de proporciones de artículos defectuosos de por lo menos 15% a favor de la máquina 1 es de 0.2357.

varianza

Ejemplos:

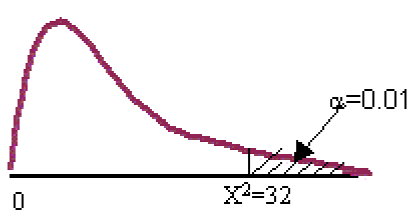
1. Suponga que los tiempos requeridos por un cierto autobús para alcanzar un de sus destinos en una ciudad grande forman una distribución normal con una desviación estándar σ =1 minuto. Si se elige al azar una muestra de 17 tiempos, encuentre la probabilidad de que la varianza muestral sea mayor que 2.

   Solución:

   Primero se encontrará el valor de ji-cuadrada correspondiente a s^2=2 como sigue:

   $$X^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(17-1)(2)}{(1)^2} = 32$$

   El valor de 32 se busca adentro de la tabla en el renglón de 16 grados de libertad y se encuentra que a este valor le corresponde un área a la derecha de 0.01. En consecuencia, el valor de la probabilidad es $P(s^2>2)$



2. Encuentre la probabilidad de que una muestra aleatoria de 25 observaciones, de una población normal con varianza $\sigma^2 = 6$

, tenga una varianza muestral:

a. Mayor que 9.1
b. Entre 3.462 y 10.745

Solución.

a. Primero se procederá a calcular el valor de la ji-cuadrada:

$$X^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(25-1)(9.1)}{6} = 36.4$$

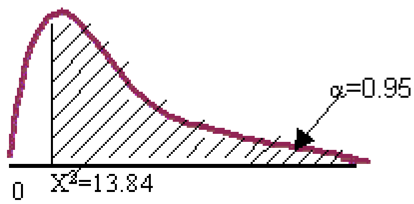
Al buscar este número en el renglón de 24 grados de libertad nos da un área a la derecha de 0.05. Por lo que la P(s^2 >9.1) = 0.05

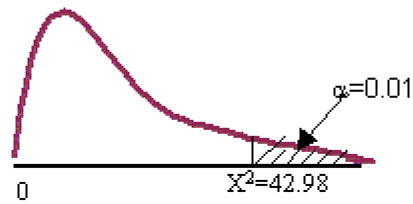
1. Se calcularán dos valores de ji-cuadrada:

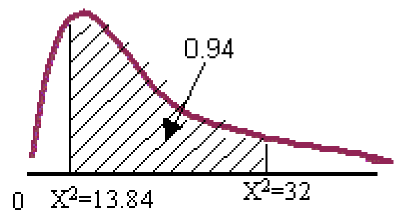
$$X^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(25-1)(3.462)}{6} = 13.847 \quad \text{y} \quad X^2 = \frac{(25-1)(10.745)}{6} = 42.98$$

Aquí se tienen que buscar los dos valores en el renglón de 24 grados de libertad. Al buscar el valor de 13.846 se encuentra un área a la derecha de 0.95. El valor de 42.98 da un área a la derecha de 0.01. Como se está pidiendo la probabilidad entre dos valores se resta el área de 0.95 menos 0.01 quedando 0.94.

Por lo tanto la P(3.462 $\leq s^2 \leq$ 10.745) = 0.94







DIFEREN PROPO

En una muestra aleatoria de 85 soportes para el cigüeñal de un motor de automóvil, 10 tienen un terminado que es más rugoso de los que las especificaciones permiten. Supóngase que se hace una modificación al proceso de acabado de la superficie y que, de manera subsecuente, se toma una segunda muestra de 85 ejes. El número de ejes defectuosos en esta segunda muestra es de 8. Obtengase un intervalo de confianza del 95% para la diferencia en la proporción de los soportes defectuosos producidos por ambos procesos y pruebe la hipótesis de que la proporción de soportes defectuosos producidos por ambos procesos es la misma.

Solución.

De lo observado en las muestras se obtiene que $\hat{p}_1 = \frac{10}{85} = 0.1176$ y $\hat{p}_2 = \frac{8}{85} = 0.09411$. El interés es la diferencia en la proporción de los soportes defectuosos entre $P_1 - P_2$:

$$H_0 : P_1 = P_2 \ vs \ H_a : P_1 \neq P_2$$

Un intervalo de confianza del 95% para la diferencia en la resistencia a la tensión promedio es:

$$(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} \le P_1 - P_2 \le (\hat{p}_1 - \hat{p}_2) - z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

$$(0.1176 - 0.09411) - 1.96\sqrt{\frac{(0.1176)(0.8824)}{85} + \frac{(0.09411)(0.90589)}{85}} \le P_1 - P_2$$

$$\le (0.1176 - 0.09411) + 1.96\sqrt{\frac{(0.1176)(0.8824)}{85} + \frac{(0.09411)(0.90589)}{85}}$$

$$-0.06893 \le P_1 - P_2 \le 0.11598$$

Este intervalo de confianza incluye al cero, así que, con base en los datos muestrales, parece poco probable que los cambios hechos en el proceso de acabado de la superficie hayan reducido el número de soportes defectuosos para cigüeñal producidos por el proceso.

Si se utiliza el estadístico presentado en (1), se encuentra:

$$Z_C = \frac{(0.1176 - 0.09411) - 0}{\sqrt{\frac{(0.1176)(0.8824)}{85} + \frac{(0.09411)(0.90589)}{85}}} = 0.49812 > z_{0.025} = 1.96$$

Rechazandose tambien la hipótesis nula, por lo tanto se concluye que los cambios hechos en el proceso de acabado de la superficie no han reducido el número de soportes defectuosos para cigüeñal producidos por el proceso.