









Inferencias Basadas En Proporciones

Estadística

Alfredo Rivera Torres

Contenido

| | | | |
|--|--|---|-----------------|
|  | Introducción | | |
|  | Inferencias Basadas en Proporciones para una Población |  | Práctica |
|  | Pruebas de Hipótesis para Proporciones en una Población |  | Práctica |
|  | Inferencias para la Diferencia Entre Dos Proporciones |  | Práctica |
|  | Ejercicios de Repaso | | |

Introducción



Los métodos para hacer inferencias (intervalos de confianza y pruebas de hipótesis) que estudiaremos ahora son similares a los que llevamos a cabo con la media, pero aplicados a otro parámetro: proporciones en poblaciones (por cientos o probabilidades).

Proporciones, por cientos y probabilidades significan lo mismo. Por ejemplo, si 33 de 100 personas prefieren un programa de noticias en particular, entonces esa proporción es igual a decir que el 33% de los televidentes en esa muestra prefiere ese programa de noticias, y que la probabilidad de seleccionar a un televidente de la muestra que prefiera ese noticiero es 0.33.

Los datos que vamos a obtener de las muestras, corresponden a una de dos categorías posibles: éxito o fracaso, ganar o perder, a favor o en contra, etc. Las proporciones, por cientos o probabilidades se calculan mediante la razón de la cantidad de datos en una de esas dos categorías con respecto a la cantidad total de datos de la muestra.

Quiere decir que vamos a trabajar con datos discretos, y si además son independientes, podemos basar los métodos de estimación en la distribución binomial, la cual cumple con esos requisitos. Nos proponemos hacer ahora inferencias con respecto a una proporción “p” de una población, llamado parámetro binomial “p”, usando la proporción \hat{p} (se lee: “p” sombrero) de una muestra como estimación de punto para “p”.



Inferencias Basadas En Proporciones Para Una Población



El parámetro binomial "p" de una población representa la probabilidad de éxito en cada intento de un experimento binomial. La probabilidad binomial de una muestra, \hat{p} , se define como: $\hat{p} = \frac{x}{n}$, donde "x" representa la cantidad de éxitos en una muestra que consiste de "n" intentos.

Sabemos que la variable binomial "x" tiene media $\mu_x = np$ y desviación estándar $\sigma_x = \sqrt{npq}$, donde "q" es la probabilidad de fracaso en cada intento, o sea, $q = 1 - p$. La distribución de "x" es aproximadamente normal cuando $n > 20$, o "n" es menos que el 10% de la población, y np y nq son ambos mayores que 5. Como nuestro deseo es hacer inferencias con respecto a "p", dividimos entre "n" a la variable "x" y a los parámetros μ_x y σ_x para obtener:

$$\hat{p} = \frac{x}{n}, \mu_{\hat{p}} = p \text{ y } \sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}.$$

Por lo tanto, si una muestra de tamaño "n" es tomada al azar de una población grande con $p = P(\text{éxito})$, entonces la distribución de \hat{p} tiene:

- 1) una media $\mu_{\hat{p}}$ igual a "p",
- 2) un error estándar $\sigma_{\hat{p}}$ igual a $\sqrt{(pq) / n}$, y
- 3) una distribución aproximadamente normal.



Cuando halla que hacer inferencias con respecto al parámetro binomial “p” de una población, P(éxito), basaremos nuestras estimaciones en la estadística no sesgada \hat{p} de una muestra. La estimación de punto para el centro del intervalo de confianza para “p” será \hat{p} . El error máximo en la estimación de “p”, así como los extremos del intervalo de confianza se consiguen mediante:

$$E = z(\alpha / 2) \cdot \sqrt{\frac{\hat{p} \hat{q}}{n}}, \text{ donde } \hat{q} = 1 - \hat{p}.$$

Extremo menor del intervalo de confianza : $\hat{p} - E$

Extremo mayor del intervalo de confianza : $\hat{p} + E$.

Como estamos estimando “p” y por lo tanto tampoco conocemos “q”, tenemos que sustituirlos con \hat{p} y \hat{q} para calcular el error estándar en la fórmula de “E”.

Ejemplo: Se lleva a cabo una encuesta para estimar la proporción de hogares en los que hay al menos una computadora. De 350 personas encuestadas, 75 tienen por lo menos una en sus casas. Determina la estimación de punto para la proporción “p” de la población que tiene computadoras y halla el intervalo del 95% de confianza para “p”.

La estimación de punto para “p” es $\hat{p} = 75/350 = 3/14$. Como $np = 350(3/14) = 75$ y $nq = 350(11/14) = 275$ son ambos mayores que 5, entonces la distribución \hat{p} es normal y podemos utilizar la estadística “z” para calcular “E”.



$$E = z(\alpha/2) \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}, \text{ donde } \hat{p} = \frac{3}{14}, \hat{q} = 1 - \hat{p} = 1 - \frac{3}{14} = \frac{11}{14}.$$

Como nos piden un intervalo con un 95% de confianza (o sea, $1 - \alpha = 0.95$), entonces $z(\alpha/2) = z(.05 / 2) = z(.025)$. El valor de $z(.025)$ en la tabla de valores de la distribución normal estándar es 1.96. La fórmula para “E” se transforma en:

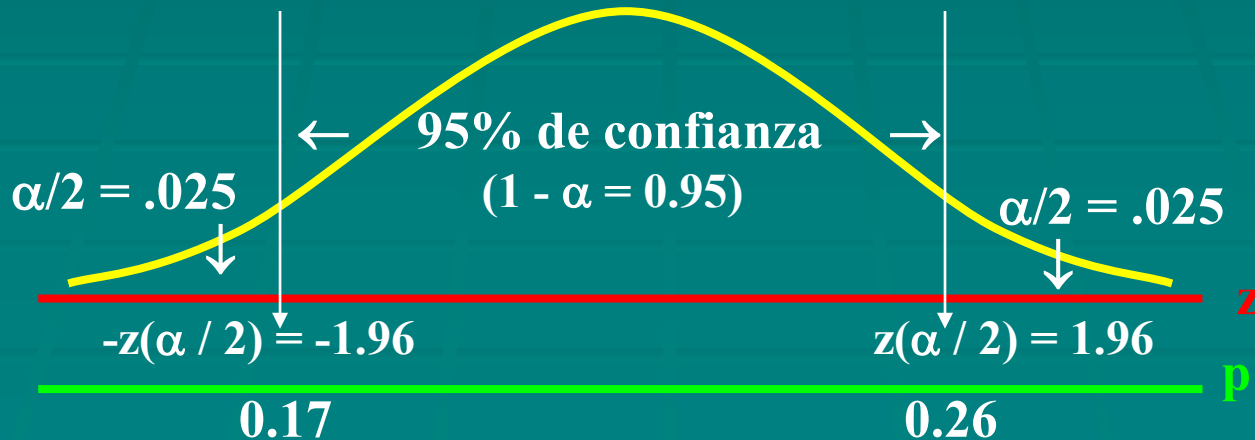
☞ Ver tabla de valores críticos de z

$$E = 1.96 \cdot \sqrt{\frac{\left(\frac{3}{14}\right)\left(\frac{11}{14}\right)}{350}} = 0.043$$

Los extremos del intervalo de confianza son:

Extremo menor del intervalo de confianza : $\hat{p} - E = \frac{3}{14} - 0.043 = 0.17$.

Extremo mayor del intervalo de confianza : $\hat{p} + E = \frac{3}{14} + 0.043 = 0.26$.



Vamos a hallar el intervalo de confianza para el ejemplo anterior usando una calculadora TI-83.



Ejemplo: Se lleva a cabo una encuesta para estimar la proporción de hogares en los que hay al menos una computadora. De 350 personas encuestadas, 75 tienen por lo menos una en sus casas. Determina el intervalo del 95% de confianza para “p”.

STAT TEST 1-PropZInt

x: 75

n: 350

C-Level .95

Calculate

```
1-PropZInt
x: 75
n: 350
C-Level: .95
Calculate
```

Oprime para comparar el IC dado por la calculadora con el que calculamos anteriormente

```
1-PropZInt
(.1713, .25727)
P=.2142857143
n=350
```



I. Halla \hat{p} en cada ejercicio.

1) $x = 14$, $n = 25$, $\hat{p} = \underline{\quad}$ 2) $x = 33$, $n = 205$, $\hat{p} = \underline{\quad}$ 3) 90 éxitos 125 intentos. $\hat{p} = \underline{\quad}$

II. Halla α , $\alpha/2$, y $z(\alpha/2)$ para el intervalo de confianza que corresponde al nivel de confianza $1 - \alpha$ dado en cada ejercicio.

4) $1 - \alpha = 0.75$, $\alpha = \underline{\quad}$, $\alpha/2 = \underline{\quad}$, $z(\alpha/2) = \underline{\quad}$

5) $1 - \alpha = 0.80$, $\alpha = \underline{\quad}$, $\alpha/2 = \underline{\quad}$, $z(\alpha/2) = \underline{\quad}$

6) $1 - \alpha = 0.90$, $\alpha = \underline{\quad}$, $\alpha/2 = \underline{\quad}$, $z(\alpha/2) = \underline{\quad}$

7) $1 - \alpha = 0.95$, $\alpha = \underline{\quad}$, $\alpha/2 = \underline{\quad}$, $z(\alpha/2) = \underline{\quad}$

Ver tabla de valores críticos de z

$$E = z(\alpha/2) \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}, \text{ donde } \hat{q} = 1 - \hat{p}.$$

Extremo menor del intervalo de confianza : $\hat{p} - E$

Extremo mayor del intervalo de confianza : $\hat{p} + E$.

III. Halla el intervalo de confianza que se pide en cada situación.

8) Una muestra al azar de 200 estudiantes del RUM, de los cuales 113 tienen beca. Halla el intervalo del 95% de confianza para la verdadera proporción “p” de todos los estudiantes que tienen beca en esa universidad.

9) La Compañía de Turismo hizo una visita a 31 Restaurantes, de los cuales el 30% incluye en su menú jugo de frutas naturales y vegetales en las tres comidas del día. Determina el intervalo con un 95% de confianza para la verdadera proporción de la población que cumple con ese requisito.

Ver clave

Ver continuación de ejercicios

10) Una muestra aleatoria de 90 mendigos en San Juan reveló que el 20% de ellos tenían estudios universitarios. Encuentra el intervalo del 99% de confianza de la verdadera proporción de mendigos con estudios universitarios de la población de San Juan.



11) Un periódico informa que de 300 policías entrevistados, 126 de ellos dijeron haber pasado la prueba de tiro al blanco. Encuentra el intervalo del 95% de confianza para la verdadera proporción de policías que han pasado esa prueba.

12) Los primeros 4,000 resultados de las elecciones generales le dan al candidato a la gobernación, Pedro Luis Acevedo Berrios el 52% de los votos. ¿deberá sentirse victorioso con un 99% de confianza?

Ver tabla de valores críticos de z

$$E = z(\alpha/2) \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}, \text{ donde } \hat{q} = 1 - \hat{p}.$$

Extremo menor del intervalo de confianza : $\hat{p} - E$

Extremo mayor del intervalo de confianza : $\hat{p} + E$.

Ver clave



Pruebas De Hipótesis Para Proporciones En Una Población

La estadística que vamos a utilizar para el parámetro binomial "p" que ha de ser utilizado en una prueba de hipótesis es:

$$z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{p \cdot q}{n}}}, \text{ donde } \hat{p} = \frac{x}{n} \text{ y } q = 1 - p.$$

Ésta tiene una distribución normal cuando la hipótesis nula es cierta, $n > 20$, np y nq son ambos mayores que 5, y se cumplen las condiciones para las probabilidades binomiales.

Ejemplo: Un estudio reporta que el número de visitas al cine durante el fin de semana es 61% más que el resto de la semana. De una muestra aleatoria de 70 clientes se encontró que 47 de ellos (67%) fue al cine en el fin de semana pasado. Con un nivel de significado de 0.05, ¿podemos demostrar que más del 61% de las personas que visitan el cine lo hacen en los fines de semana?

Podemos aproximar la distribución de "p" usando la distribución normal estándar "z", porque $n = 70 > 20$, $np = 70(.61) = 42.7 > 5$, $nq = 70(.39) = 27.3 > 5$, y hay solamente dos posibles resultados para cada individuo: fin de semana o el resto de la semana. La estimación de punto para "p" es $\hat{p} = 47/70$.



Paso # 1 :

$H_a : p > 0.61$ (61%)

$H_0 : p \leq 0.61$

Paso # 2:

Nivel de significado:

$\alpha = 0.05$

Paso # 3 : Criterio de decisión.

Rechazo H_0 si $z_{47/70} > 1.65$

☞ Ver tabla de valores críticos de z

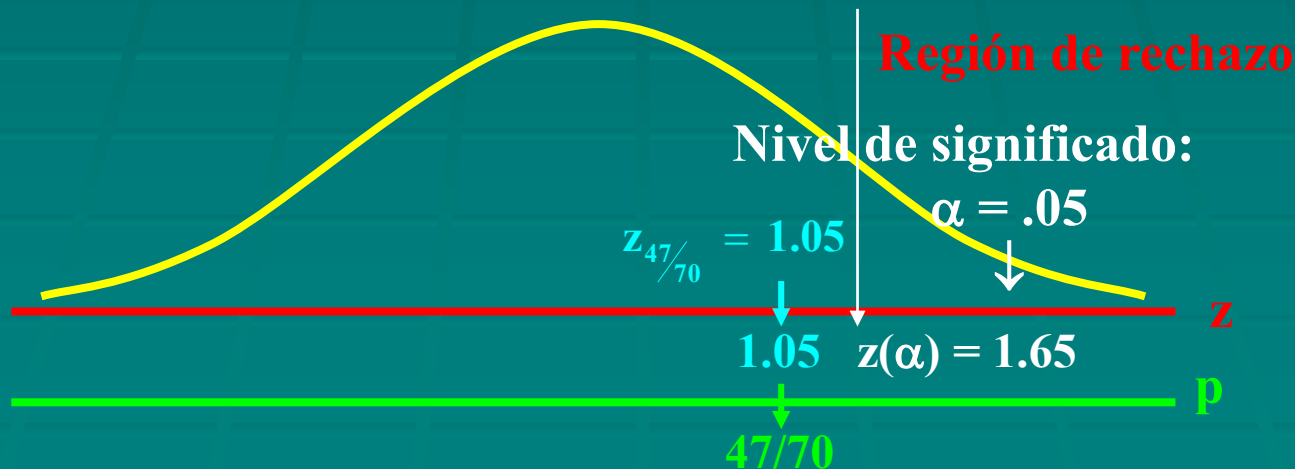
Paso # 4 : Cálculos.

$$z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \square z_{47/70} = \frac{\frac{47}{70} - 0.61}{\sqrt{\frac{(.61)(.39)}{70}}} = 1.05$$

Paso # 5 : Decisión.

Como $z_{47/70} = 1.05$ es menor que 1.65, no rechazamos H_0 .

No tenemos suficiente evidencia en contra de H_0 .





Vamos a hacer esa prueba utilizando la calculadora TI-83. Ésta reporta sus resultados usando el método del “p - value”.

Recuerda que en esta prueba, el valor de “p” es el de H_0 , o sea, $p_0 = 0.61$, $x = 47$ y $n = 70$. La hipótesis alterna es $H_a > 0.61$.

STAT TEST 1-PropZTest

1-PropZTest
 P0: .61
 x: 47
 n: 70
 PROP≠P0 <P0 >P0
 Calculate Draw

p 0:
 x:
 n:
 prop
 Calculate

Oprime para comparar la decisión a la cual llegamos con el uso de la calculadora con la que llegamos con la prueba tradicional.

```

1-PropZTest
PROP>.61
z=1.0537127
P=.1460072619
P̂=.6714285714
n=70
  
```

Como $p > \alpha$, no rechazamos H_0 .



I. Escribe la H_a y la H_o que debes usar en la prueba de cada situación.

1) Menos del 40% de todos los graduados de pedagogía en Puerto Rico consiguen trabajo durante el primer semestre después de graduarse.

H_a : _____ H_o : _____

2) La probabilidad de que llueva esta tarde es más de 50%.

H_a : _____ H_o : _____

3) No menos de una tercera parte de los diabéticos están interesados en seguir una dieta.

H_a : _____ H_o : _____

4) Por lo menos el 30% de las parejas que se casan se divorcian al cabo de un año.

H_a : _____ H_o : _____

II. Determina el criterio de decisión que debes utilizar en cada ejercicio cuando usas “z” como la estadística de prueba.

5) $H_o: p \leq 0.4$, $H_a: p > 0.4$, $\alpha = .02$, criterio: _____

6) $H_o: p = 0.3$, $H_a: p \neq 0.3$, $\alpha = .02$, criterio: _____

7) $H_o: p \geq 0.5$, $H_a: p < 0.5$, $\alpha = .05$, criterio: _____

8) $H_o: p = 0.8$, $H_a: p \neq 0.8$, $\alpha = .25$, criterio: _____

Ver tabla de valores críticos de z

$$z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

Ver clave

Ver continuación de ejercicios



III. Para las siguientes hipótesis, decide si hay que usar una prueba de uno o de dos extremos, y en el caso que sea de un solo extremo, indica en cuál extremo está la región de rechazo.

9) Menos del 3% de las personas que cepillan sus dientes tres veces al día tienen caries.

10) Más del 30% de los maestros de la escuela llega tarde a su salón de clases.

11) Los estudiantes de medicina del Dr. Selástraga llevan un promedio de 60% en su clase. Decide sorprenderlos con exámenes inesperados en cirugía de corazón abierto con el propósito de cambiar dicho por ciento de aprovechamiento.

$$z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

IV. Realiza la prueba de hipótesis que se requiere.

12) La tasa anual de divorcios en Puerto Rico es 50%. El nuevo director del Departamento de la Familia afirma que de acuerdo a un nuevo plan, han disminuido la tasa de divorcios, ya que en este año, de 500 matrimonios nuevos, solamente se divorciaron 230. Prueba esa proporción con un nivel de significado de 0.02. ¿Apoyan las estadísticas tal afirmación?

13) La profesora Enid M. Atica opina que el 70% de los números que se escriben al azar son pares. De acuerdo a la siguiente tabla de números aleatorios, demuestra lo que afirma Enid con un nivel de significado de 0.05.

| | | | | | |
|---|---|---|---|---|---|
| 3 | 1 | 5 | 8 | 5 | 3 |
| 8 | 4 | 4 | 1 | 5 | 2 |
| 6 | 4 | 3 | 2 | 6 | 8 |
| 0 | 3 | 1 | 1 | 0 | 7 |
| 7 | 8 | 5 | 4 | 4 | 0 |
| 2 | 4 | 3 | 3 | 0 | 3 |

Ver tabla de valores críticos de z

Ver clave

Para continuar



14) Un oficial del Departamento de Asuntos al Consumidor afirma que el 70% de las garantías termina debido al límite máximo de millas, en lugar de la alternativa del tiempo máximo estipulado. Para probar dicha proporción, seleccionamos a 60 clientes a los que se le venció la garantía de su auto. Si a 38 de ellos se les venció por alcanzar el máximo de millas, ¿es esa proporción evidencia de lo que proclama la compañía, con un nivel de significado de 0.05?

$$z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{p q}{n}}}$$

[🔗 Ver tabla de valores críticos de z](#)

[🔗 Ver clave](#)



Inferencias Para La Diferencia Entre Proporciones Basadas En Dos Muestras Independientes

Deseamos comparar ahora las proporciones, por cientos o probabilidades para dos poblaciones usando la diferencia entre las proporciones de dos muestras independientes $\hat{p}_1 - \hat{p}_2$. Esas diferencias observadas para las proporciones de las muestras, $\hat{p}_1 - \hat{p}_2$, provienen de una distribución de muestreo, cuyas características son las siguientes:



Si muestras independientes de tamaños n_1 y n_2 son obtenidas al azar de poblaciones grandes con proporciones $p_1 = P_1(\text{éxito})$ y $p_2 = P_2(\text{éxito})$, respectivamente, entonces la distribución de muestreo para $\hat{p}_1 - \hat{p}_2$ tiene las siguientes características:

1) una media igual a la diferencia de las proporciones de las poblaciones, o sea, $\mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2$.

2) un error estándar $\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$

3) una distribución aproximadamente normal si n_1 y n_2 son grandes.

En la práctica, la normalidad de la distribución de las diferencias entre proporciones queda garantizada si cada muestra tiene $n > 20$, los productos $n_1 p_1$, $n_1 q_1$, $n_2 p_2$, y $n_2 q_2$ son cada uno mayores que 5, y las muestras son menos del 10% de la población.



Si p_1 y p_2 son desconocidos, como suele suceder, entonces los productos n_1p_1 , n_1q_1 , n_2p_2 y n_2q_2 serán estimados usando en sus lugares $n_1\hat{p}_1$, $n_1\hat{q}_1$, $n_2\hat{p}_2$ y $n_2\hat{q}_2$.

Intervalo de confianza para $p_1 - p_2$

Para estimar la diferencia entre dos proporciones $p_1 - p_2$, vamos a utilizar la estadística no sesgada $\hat{p}_1 - \hat{p}_2$. El valor de $\hat{p}_1 - \hat{p}_2$ va a ser la estimación de punto del intervalo de confianza, cuyos extremos

son $(\hat{p}_1 - \hat{p}_2) - E$ y $(\hat{p}_1 - \hat{p}_2) + E$, donde $E = z(\alpha/2) \cdot \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$.

Ejemplo: En una encuesta a 300 personas de Aguadilla, 128 prefirieron el jabón Love sobre las demás marcas, mientras que en Bayamón lo prefirieron 149 de 400. Halla el intervalo con el 98% de confianza para la diferencia en las dos proporciones.

La distribución de muestreo de $\hat{p}_A - \hat{p}_B$ es aproximadamente normal, ya que el tamaño de cada muestra es mayor que 20, y además, $n_A \hat{p}_A$, $n_A \hat{q}_A$, $n_B \hat{p}_B$ y $n_B \hat{q}_B$ son cada uno mayores que 5. Como $n_A = 300$, $x_A = 128$, $n_B = 400$


y $x_B = 149$, entonces $\hat{p}_A = \frac{128}{300} = \frac{32}{75}$, $\hat{q}_A = 1 - \frac{32}{75} = \frac{43}{75}$, $\hat{p}_B = \frac{149}{400}$ y

$$\hat{q}_B = 1 - \frac{149}{400} = \frac{251}{400}.$$

☞ Ver tabla de valores críticos de z

La estimación de punto para $(p_A - p_B)$ es $\frac{32}{75} - \frac{149}{400} = \frac{13}{240}$.

Como $1 - \alpha = 98\% = 0.98$, entonces $\alpha = 0.02$ y $\alpha / 2 = 0.01$. El valor para $z(\alpha / 2) = z(.01) = 2.33$. (Vea tabla de valores de “z” de la curva normal estándar.)



$$E = z(\alpha / 2) \cdot \sqrt{\frac{\hat{p}_A \hat{q}_A}{n_A} + \frac{\hat{p}_B \hat{q}_B}{n_B}} = 2.33 \cdot \sqrt{\frac{\frac{32}{75} \cdot \frac{43}{75}}{300} + \frac{\frac{149}{400} \cdot \frac{251}{400}}{400}} = 0.087$$

Extremos del intervalo de confianza :

extremo menor : $(\hat{p}_A - \hat{p}_B) - E = \frac{13}{240} - 0.087 = -0.033$

extremo mayor : $(\hat{p}_A - \hat{p}_B) + E = \frac{13}{240} + 0.087 = 0.141$

Puedes hallar ese intervalo de confianza usando un programa de computadora o algunas calculadoras gráficas.



Vamos a hallar el intervalo de confianza del ejemplo anterior usando una calculadora TI-83.

Ejemplo: En una encuesta a 300 personas de Aguadilla, 128 prefirieron el jabón Love sobre las demás marcas, mientras que en Bayamón lo prefirieron 149 de 400. Halla el intervalo con el 98% de confianza para la diferencia en las dos proporciones.

STAT

TEST

2-PropZInt

x 1:

n 1:

x 2:

n2:

C-Level:

Calculate:

```
2-PropZInt
x1: 128
n1: 300
x2: 149
n2: 400
C-Level: .98
Calculate
```

Oprime para comparar el IC dado por la calculadora con el que calculamos anteriormente

```
2-PropZInt
(-.0329, .1412)
P1 = .426666666667
P2 = .3725
n1 = 300
n2 = 400
```



Pruebas de hipótesis para $p_1 - p_2$ usando dos muestras independientes

Las hipótesis nulas de las pruebas que llevaremos a cabo en esta sección serán del tipo en las que no hay diferencia entre las proporciones correspondientes a dos poblaciones: $p_1 - p_2 = 0$, o sea, $p_1 = p_2$.

El valor crítico de "z" para $\hat{p}_1 - \hat{p}_2$ en los cálculos es:

$$Z_{\hat{p}_1 - \hat{p}_2} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{pq \left[\left(\frac{1}{n_1} \right) + \left(\frac{1}{n_2} \right) \right]}}$$

A continuación presentamos algunas consideraciones al momento de aplicar esa fórmula al hacer los cálculos.

Notas :



▷ El numerador de $z_{\hat{p}_1 - \hat{p}_2}$, es en realidad $(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)$. Aparece solamente $\hat{p}_1 - \hat{p}_2$ porque $p_1 - p_2 = 0$.



Ejemplo: En una revisión de los archivos de defunciones en un hospital, se encontró que de 50 personas no fumadoras, 6 murieron de cáncer pulmonar mientras que de 60 fumadores, 15 fallecieron a consecuencia de ese tipo de cáncer. Utiliza una prueba de hipótesis con un nivel de significado de .05 para determinar si la diferencia en las proporciones es significativa.

Para esta situación, $\hat{p}_f = \frac{x_f}{n_f} = \frac{15}{60} = 0.25,$

$\hat{p}_n = \frac{x_n}{n_n} = \frac{6}{50} = 0.12,$ $\hat{p}_p = \frac{x_f + x_n}{n_f + n_n} = \frac{15 + 6}{60 + 50} = \frac{21}{110}$

🔍 Ver tabla de valores críticos de z

Primer paso :

$H_a : p_f - p_n > 0,$ o sea, $p_f > p_n$

$H_o : p_f - p_n \leq 0,$

Paso # 2.

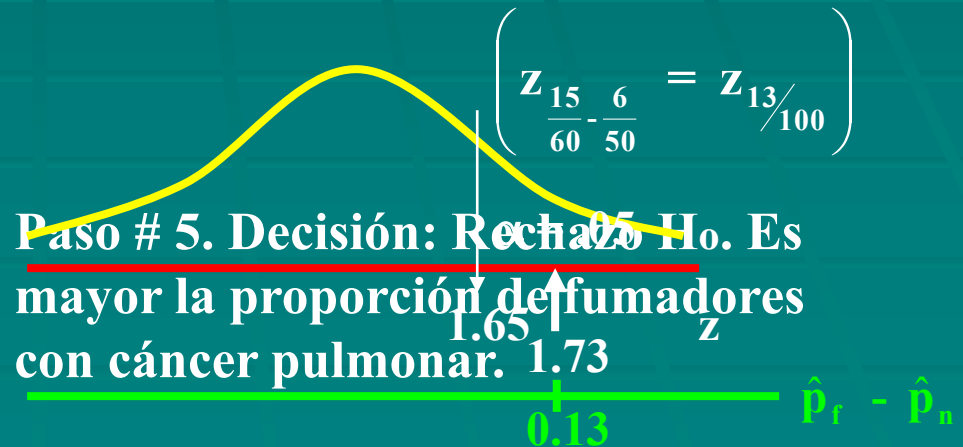
Nivel de significado:
0.05 en un extremo

Paso # 3.

Criterio : rechazo H_o
si $z_{13/100} > 1.65.$

Paso # 4. Cálculos :

$$z_{13/100} = \frac{13/100}{\sqrt{\frac{21}{110} \cdot \frac{89}{110} \left(\frac{1}{60} + \frac{1}{50} \right)}} = 1.73$$





Vamos a hacer la pasada prueba mediante la calculadora TI-83. Recuerda que ésta presenta los resultados usando el método del “p-value”.

Ejemplo: En una revisión de los archivos de defunciones en un hospital, se encontró que de 50 personas no fumadoras, 6 murieron de cáncer pulmonar mientras que de 60 fumadores, 15 fallecieron a consecuencia de ese tipo de cáncer. Utiliza una prueba de hipótesis con un nivel de significado de .05 para determinar si la diferencia en las proporciones es significativa.

STAT TEST 2-PropZTest

x 1: 15
 x 1: 60
 x 2: 6
 n2: 50
 p1:
 Calculate:

2-PropZTest
 x1: 15
 n1: 60
 x2: 6
 n2: 50
 P1: ≠ P2 < P2
 Calculate Draw

Oprime para comparar la decisión a la cual llegamos con el uso de la calculadora con la que llegamos con la prueba tradicional.

Como $p < \alpha$, rechazo H_0 .

2-PropZTest
 P1 > P2
 z=1.727410963
 P=.0420469036
 P1=.25
 P2=.12
 P=.1909090909



I. Halla el intervalo de confianza que se pide en cada situación.

- 1) De 40 fanáticos entrevistados en San Germán, 22 dijeron que seguían los juegos de sus Atléticos. En otra muestra al azar de 40 personas de Quebradillas, 26 dijeron que seguían a sus Piratas en el baloncesto. Utiliza un intervalo con 98% de confianza para estimar la diferencia en la proporción de los que siguen a su equipo en esos dos pueblos.

Ver tabla de valores críticos de z

- 2) La proporción de galletas rotas producidas por dos máquinas es comparada en la siguiente tabla:

| Máquina | Tamaño de la muestra | # de galletas rotas |
|---------|----------------------|---------------------|
| # 1 | 150 | 12 |
| # 2 | 150 | 6 |

Determina el intervalo del 90% de confianza para $p_1 - p_2$.

II. Escribe la hipótesis alterna H_a y la nula H_0 que deben utilizarse en cada proposición.

- 3) Hay diferencia en la proporción de hombres y mujeres pacientes de cáncer pulmonar.

H_a : _____ H_0 : _____

- 4) El por ciento de muchachas que van a universidad es mayor que el de los varones.

H_a : _____ H_0 : _____

- 5) El por ciento de jóvenes que estudian en la zona rural es mayor que el por ciento de jóvenes que estudian de la zona urbana. H_a : _____ H_0 : _____

Ver clave

Ver continuación de ejercicios

Extremos del I.C.:

$$(\hat{p}_1 - \hat{p}_2) - E \text{ y } (\hat{p}_1 - \hat{p}_2) + E, E = z(\alpha/2) \cdot \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

III. Determina el criterio de decisión que debes utilizar en cada ejercicio cuando usas “z” como la estadística de prueba.

6) $H_0: p_1 \geq p_2$, $H_a: p_1 < p_2$, $\alpha = .01$, criterio: _____

7) $H_0: p_1 \leq p_2$, $H_a: p_1 > p_2$, $\alpha = .02$, criterio: _____

8) $H_0: p_1 - p_2 = 0$, $H_a: p_1 - p_2 \neq 0$, $\alpha = .01$, criterio: _____

9) $H_0: p_1 - p_2 = 0$, $H_a: p_1 - p_2 \neq 0$, $\alpha = .02$, criterio: _____

$$z_{\hat{p}_1 - \hat{p}_2} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{pq [(1/n_1) + (1/n_2)]}}$$

IV. Realiza una prueba de hipótesis para la diferencia de proporciones en cada ejercicio.

10) Se sospecha que la proporción de hombres que comen tres veces al día es mayor que para las mujeres. De 500 hombres escogidos al azar, 432 acostumbran hacer las tres comidas al día, mientras que de 500 mujeres entrevistadas, 401 comen tres veces por día. Determina si la diferencia en proporciones es significativa. Utiliza $\alpha = 0.01$.

11) En los procedimientos médicos en los que se utiliza anestesia, hay la posibilidad de que el paciente baje la presión. Se lleva a cabo un experimento clínico en el que a 19 pacientes se le suministró anestesia espinal, de los cuales 6 bajaron la presión. A otros 25 se les aplicó anestesia general, de los cuales 18 bajaron la presión. De acuerdo a esos datos, ¿podemos afirmar que la anestesia general baja la presión significativamente más que la espinal? Utiliza un nivel de significado $\alpha = .01$.

 Ver clave

 Ver tabla de valores críticos de z

Ejercicios de Repaso: Inferencias Basadas en Proporciones



I. Contesta Cierto o Falso de acuerdo a \hat{p} .

- 1) El valor de \hat{p} es la proporción de éxitos en una población.
- 2) La estimación de punto para "p" de una población es \hat{p} de una muestra.
- 3) \hat{p} es un parámetro.
- 4) La distribución de muestreo de una población binómica grande cuya proporción es "p", tiene una media $\mu_{\hat{p}} = p$ si el tamaño de las muestras es $n > 20$, $np > 5$ y $nq > 5$.
- 5) El error estándar de la distribución de muestreo para "p" es $\sigma_{\hat{p}} = \sqrt{npq}$.
- 6) El centro del intervalo de confianza para "p" es \hat{p} .

II. Dado el tamaño "n" de una muestra, la cantidad de resultados favorables "x" y el grado de confianza $1 - \alpha$, completa la tabla con el error máximo (E) en la estimación de "p" y su intervalo de confianza (IC).

| | n | x | \hat{p} | $1 - \alpha$ | E | IC |
|----|-------|-------|-----------|--------------|---|----|
| 7) | 500 | 120 | | 0.98 | | |
| 8) | 2,400 | 1,125 | | 0.90 | | |

$$E = z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}$$

[Ver clave](#)

[Ver continuación de ejercicios](#)

[Ver tabla de valores críticos de z](#)



II. Dado el tamaño “n” de una muestra, el por ciento de resultados favorables (%) y el grado de confianza $1 - \alpha$, completa la tabla con el error máximo (E) en la estimación de “p” y su intervalo de confianza.

$$E = z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}$$

| | n | % | \hat{p} | $1 - \alpha$ | E | Intervalo de confianza |
|-----|------|-----|-----------|--------------|---|------------------------|
| 9) | 1000 | 18% | | 99% | | |
| 10) | 144 | 45% | | 95% | | |

III. Halla el criterio de decisión en cada ejercicio de acuerdo a los datos de la tabla.

| | H_a | H_0 | α | Localización zona de rechazo (extremo derecho, izquierdo o ambos) | Criterio de decisión |
|-----|---------------|-------------|----------|--|----------------------|
| 11) | $p > 1/5$ | | .025 | | |
| 12) | | $p \geq .6$ | 0.02 | | |
| 13) | $p \neq 30\%$ | | 0.01 | | |

IV. Completa la tabla según las estadísticas dadas de dos muestras independientes.

| | n1 | n2 | x1 | x2 | \hat{p}_1 | \hat{p}_2 | $1 - \alpha$ | $z(\alpha/2)$ | Intervalo de confianza |
|-----|-----|-----|-----|-----|-------------|-------------|--------------|---------------|------------------------|
| 14) | 500 | 500 | 220 | 190 | | | 0.90 | | |
| 15) | 70 | 55 | 60 | 44 | | | .75 | | |

$$E = z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}_1 \cdot \hat{q}_1}{n_1} + \frac{\hat{p}_2 \cdot \hat{q}_2}{n_2}}$$

[Ver clave](#)

[Ver continuación de ejercicios](#)

[Ver tabla de valores críticos de z](#)



V. Realiza una prueba de hipótesis para cada problema.

16) Se afirma que de 600 personas entrevistadas que pagan hipotecas, el 42% se siente preocupada de poder cumplir con el pago. Se hace una encuesta similar de la que se obtiene 196 de 600 personas también preocupadas por dicho pago. Demuestra que el porcentaje es significativamente menor que 42%. Utiliza un nivel de significado de 0.05.

$$z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{p \cdot q}{n}}}$$

17) Se afirma que el 10% de los estudiantes son zurdos. Derek Sho, un estudiante zurdo, sospecha que el porcentaje de zurdos es mayor. Lleva a cabo una encuesta de la que obtiene 16 zurdos de 100 estudiantes escogidos al azar. ¿Es eso evidencia de lo que supone Derek con un nivel de significado de 0.05?

[🔑 Ver clave](#)

[🔑 Ver continuación de ejercicios](#)

[🔑 Ver tabla de valores críticos de z](#)



V. Realiza una prueba de hipótesis para cada problema.

18) “Science News” (Alcohol Distills Aggression in Large Men: Bruce Brower) reportó que los hombres corpulentos, cuando se embriagan, muestran mayor agresividad que los menos corpulentos.

La siguiente tabla muestra los resultados de un estudio sobre ese particular hecho en Puerto Rico. Prueba la hipótesis, con un nivel de significado de .05, de que la proporción de corpulentos embriagados violentos es mayor que la de los menos corpulentos embriagados violentos.

| Embriagados | n | # de violentos |
|------------------|----|----------------|
| Corpulento | 50 | 22 |
| Menos corpulento | 50 | 17 |

$$Z_{(\hat{p}_1 - \hat{p}_2)} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

19) ¿Aceptarías un segundo trabajo si ello representa menos tiempo con tu familia? De 100 hombres y 100 mujeres entrevistadas, 32% y 26% contestaron que sí, respectivamente. ¿Hay diferencia en la proporción de hombres y mujeres que respondieron que sí, con un nivel de significado de .05?

