Ejercicio 1.

El tiempo de secado de las pinturas tradicionales se distribuye normalmente condesvío estándar 9 min. Todas las pinturas del mercado tienen un tiempo medio de secado de 75 minutos o más Se diseño una aditivo químico para disminuir el tiempo medio de secado de la pintura a menos de 75 minutos. Se cree que el tiempo de secado con este aditivo también se distribuye de manera normal con $\sigma = 9$. Se decide recolectar una muestra de $X_1, \dots X_n$ tiempos de secados independientes. Sea μ el tiempo de secado medio cuando el aditivo es usado. La hipótesis son

$H_0 : \mu \geq 75$

$H_1 : \mu < 75$

1. Plantear el estadístico del test y la región de rechazo para tener un nivel $\alpha = 0.05$.
2. Si se toma una muestra de $n = 25$ tiempos de secado. Calcular la potencia del test si el tiempo medio de secado es de 73 minutos.
3. ¿Que tamaño de muestra debe tomarse si se quiere que la probabilidad de que el test logre rechazar la hipótesis nula con probabilidad al menos 0.99 en el caso que el tiempo medio verdadero para el secado con el aditivo sea de 73 minutos?

Ejercicio 2. Para corroborar que la calibración de una balanza sea adecuada, se pesará 25 veces una pesa cuyo peso verdadero se sabe que es de 10 kg. Supongamos que los resultados de diferentes pesadas de la misma pesa son independientes entre sí y que la medicion del peso de cada intento está normalmente distribuida con desviacion estandard $\sigma = 0.2kg$. Llame μ a la esperanza de una medicion del peso de la pesa de 10kg en la balanza.

1. Suponga que planteamos como hipotesis nula a $H_0 : \mu = 10kg$ y como hipotesis alternativa a $H_1 : \mu \neq 10kg$. Discuta si el planteo de las hipotesis nula y alternativa para este problema es adecuada.
2. Supongamos que establecemos la regla que indica que si el promedio de las 25 pesadas, digamos $\overline{X}_n$, satisface $\overline{X}_n \geq 10.1kg$ o $\overline{X}_n \leq 9.89kg$ entonces la balanza debe recalibrarse, de lo contrario, no. ¿Cuál es la probabilidad de que la recalibración se realice cuando en realidad no es necesaria? ¿Cuál es la probabilidad de que la recalibración se considere necesaria cuando $\mu = 10kg$?

Solución

Notar que X_i es una variable aleatoria que mide la balanza que pesa la pesa. $X_i \sim \mathcal{N}(\mu, 0.2^2)$

Entonces $\overline{X}_n \sim \mathcal{N}(\mu, 0.2^2/25) = \mathcal{N}(\mu, 1/25^2)$

O sea $\dfrac{\overline{X}_n - 10}{1/25} \sim \mathcal{N}(0, 1)$

$$\begin{cases} H_0 & \mu = 10 \text{ kg} \\ H_1 & \mu \neq 10 \text{ kg} \end{cases}$$

Consideramos que este test es adecuado para el problema ya que como nos interesa calibrar la balanza (es decir, que la balanza pese de manera precisa) queremos que no haya errores de medición ni para arriba ni para abajo.

Regla de decisión: Rechazo H_0 si $\overline{X}_n > 10.1$ o $\overline{X}_n < 9.89$

¿Cuál es la probabilidad de que la $\underbrace{\text{recalibración se realice}}_{\text{Rechazar } H_0}$ $\underbrace{\text{cuando en realidad no es necesaria}}_{\text{dado que } \mu = 10}$?

Nos están preguntando por la probabilidad del error de tipo I.

$$\begin{aligned}
&P_{\mu=10}(\text{Rechazo}H_0) \\
&P_{\mu=10}(\overline{X}_n > 10.1) + P(\overline{X}_n < 9.89) \\
&P_{\mu=10}\left(\frac{\overline{X}_n - 10}{1/25} > \frac{10.1 - 10}{1/25}\right) + P_{\mu=10}\left(\frac{\overline{X}_n - 10}{1/25} < \frac{9.89 - 10}{1/25}\right) \\
&P(Z > 2.5) + P(Z < -2.75) \approx 0.009
\end{aligned}$$

Ejercicio 3. Se ha propuesto un nuevo diseño para el sistema de frenos de un automóvil. Se sabe que, para el sistema actual, el verdadero promedio de distancia de frenado, a 65 kilómetros por hora, bajo condiciones especificadas, es de 4 metros. Se propone que el nuevo diseño se ponga en práctica sólo si los datos muestrales indican una reducción en el verdadero promedio de distancia de frenado para el nuevo diseño.

1. Defina el parámetro de interés e indique las hipótesis nula y alternativa pertinentes para establecer si debe o no poner en practica el nuevo diseño.

2. Suponga que la distancia de frenado para el nuevo sistema de una medición cualquiera tomada al azar está normalmente distribuida con desviación standard $\sigma = 10$. Denotemos con $\overline{X}_{36}$ al promedio muestral de distancia de frenado para una muestra aleatoria de 36 observaciones. Considere los siguientes tres test de hipotesis de H_0, cuyas regiones de rechazo son:

   (a) El test rechaza cuando $\overline{X}_{36} \leq 3.5$.

   (b) El test rechaza cuando $\overline{X}_{36} \leq 2.5$.

   (c) El test rechaza cuando $2.2 \leq \overline{X}_{36} \leq 2.7$.

   ¿Cuál es el nivel de significación para cada test?

3. Para cada test, establezca cuál es la probabilidad de que el nuevo diseño no se ponga en práctica cuando el verdadero promedio de distancia de frenado es en realidad 3.7 metros.

4. ¿Cuál de los tres procedimientos le parece que es el más adecuado para testear la hipotesis nula vs la hipótesis alternativa del primer inciso?

Solución

1. Definimos la variable aleatoria

   $X_i =$ 'cantidad de metros que necesitó el i-ésimo auto en frenar, si iba inicialmente a 65 kmph'

   Luego, el parámetro de interés es la media μ. Las hipótesis a testear son:

   H_0**:** $\mu = 4$,

   H_1**:** $\mu < 4$.

2. Sabemos que $X_i \sim \mathcal{N}\left(\mu, 10^2\right)$. El estadístico del test es

$$\overline{X}_{36} \sim \mathcal{N}\left(\mu, \left(\frac{10}{6}\right)^2\right)$$

   cuya estandarización bajo H_0 verifica

$$Z = \frac{\overline{X}_{36} - 4}{10/6} \sim \mathcal{N}(0,1)$$

El nivel de significación es:

$$\alpha = P(\text{rechazo } H_0 \text{ cuando es verdadera})$$

Que H_0 sea verdadera significa que $\mu = 4$. Calculemos el nivel de significación para cada región de rechazo

a) Para la primera región de rechazo su nivel de significación será

$$\alpha = P_{\mu=4}\left(\overline{X}_{36} \leq 3.5\right) = P_{\mu=4}\left(\underbrace{\frac{\overline{X}_{36} - 4}{10/6}}_{=Z\sim\mathcal{N}(0,1)} < \frac{3.5 - 4}{10/6}\right) = P_{\mu=4}(Z < -0.3) \approx 0.382$$

b) Para la segunda región el nivel de significación será

$$P_{\mu=4}\left(\overline{X}_{36} < 2.5\right) = P_{\mu=4}\left(\frac{\overline{X}_{36} - 4}{10/6} < \frac{2.5 - 4}{10/6}\right) = P_{\mu=4}(Z < -0.9) \approx 0.184$$

c) Para la tercera región de rechazo el nivel de significación será

$$\begin{aligned}
P_{\mu=4}\left(2.2 < \overline{X}_{36} < 2.7\right) &= P_{\mu=4}\left(\frac{2.2 - 4}{10/6} < \frac{\overline{X}_{36} - 4}{10/6} < \frac{2.7 - 4}{10/6}\right) \\
&= P_{\mu=4}(-1.08 < Z < -0.78) \\
&= P_{\mu=4}(Z < -0.78) - P_{\mu=4}(Z \leq -1.08) \\
&\approx 0.218 - 0.140 = 0.078
\end{aligned}$$

d) En este inciso nos piden calcular la probabilidad del error de tipo II cuando $\mu = 3.7$. Es decir

$$P(\text{no rechazo } H_0 \text{ cuando } \mu = 3.7)$$

a) Para el primer test la probabilidad del error de tipo II será

$$P_{\mu=3.7}\left(\overline{X}_{36} \geq 3.5\right) = P_{\mu=3.7}\left(\frac{\overline{X}_{36} - 3.7}{10/6} \geq \frac{3.5 - 3.7}{10/6}\right) = P_{\mu=4}(Z \geq -0.12) \approx 0.548$$

b) Para el segundo test la probabilidad del error de tipo II será

$$P_{\mu=3.7}(\overline{X}_{36} \geq 2.5) = P_{\mu=3.7}\left(\frac{\overline{X}_{36} - 3.7}{10/6} \geq \frac{2.5 - 3.7}{10/6}\right) = P_{\mu=3.7}(Z \geq -0.72) \approx 0.764$$

c) Para el tercer test la probabilidad del error de tipo II será

$$\begin{aligned} P_{\mu=3.7}(2.2 \leq \overline{X}_{36} \leq 2.7) &= 1 - P_{\mu=3.7}\left(\frac{2.2 - 3.7}{10/6} < \frac{\overline{X}_{36} - 3.7}{10/6} < \frac{2.7 - 3.7}{10/6}\right) \\ &= 1 - P(-0.9 < Z < -0.6) \\ &= 1 - (P(Z < -0.6) - P(Z \leq -0.9)) \\ &\approx 1 - (0.274 - 0.184) = 0.91 \end{aligned}$$

d) La regla c) queda descartada por no tener sentido para evaluar la hipótesis que nos interesa ya que rechaza en casos donde la distancia de frenado es significativamente menor. Entre las reglas a) y b) se prefiere la que tenga menor error de tipo I, y por eso se prefiere la regla b). Idealmente habría que tomar una muestra de mayor tamaño para poder garantizar que la potencia fuera mayor.

Ejercicio 4. Un fabricante de aspirinas esta interesado en determinar si la media del peso de las tabletas de su produccién es de 5 gramos. Con este fin ha pesado 100 tabletas elegidas al azar de un lote muy grande, resultando en un peso promedio muestral de 4.87 gramos y una desviación estándar muestral de 0.35 gramos. ¿Proporcionan estos resultados evidencia convincente como para concluir que la media del peso de cada tableta no es de 5 gramos? Establezca las hipótesis nula y alternativa para responder a esta pregunta, conduzca un test con nivel de significación asintótica 0.05.

Solución

Sea X_i el peso de una tableta extraída al azar del lote. Las variables X_i son iid y definimos $\mu = E(X_1)$ y $\sigma^2 = Var(X_1)$. Por TCL:

$$\frac{\sqrt{n}\left(\overline{X}_n - \mu\right)}{\sigma} \xrightarrow{D} N(0,1)$$

Luego, por Slutsky,

$$\frac{\sqrt{n}\left(\overline{X}_n - \mu\right)}{S} \xrightarrow{D} N(0,1)$$

Usando que el desvío estándar muestral, S, es un estimador consistente del desvío estándar poblacional. Las hipótesis a contrastar son

$$H_0 : \mu = 5$$
$$H_1 : \mu \neq 5$$

Entonces, bajo H_0:

$$\frac{\sqrt{n}\left(\overline{X}_n - 5\right)}{S} \xrightarrow{D} N(0,1)$$

Es decir que

$$P_{H_0}\left(-z_{0.975} \leq \frac{\sqrt{n}\left(\overline{X}_n - 5\right)}{S} \leq z_{0.975}\right) \to 0.95$$

$$P_{H_0}\left(\left|\frac{\sqrt{n}\left(\overline{X}_n - 5\right)}{S}\right| \leq z_{0.975}\right) \to 0.95$$

$$P_{H_0}\left(\left|\frac{\sqrt{n}\left(\overline{X}_n - 5\right)}{S}\right| > z_{0.975}\right) \to 0.05$$

Como $z_{0.975} \approx 1.96$, el test que rechaza cuando

$$\left|\frac{\sqrt{n}\left(\overline{X}_n - 5\right)}{S}\right| > 1.96$$

es un test de nivel de significación asintótica 0.05.
En el caso observado, tenemos

$$n = 100$$
$$\overline{x} = 4.87$$
$$s = 0.35$$

Luego,

$$\left|\frac{10\,(4.87 - 5)}{0.35}\right| \approx 3.7143 > 1.96$$

Por lo que hay suficiente evidencia para rechazar H_0 al nivel asintótico especificado: concluimos que le media del peso de cada tableta no es de 5 gramos.

Ejercicio 5. El gobierno de la ciudad de Buenos Aires ha realizado recientemente una vigorosa campaña de reparación de veredas. Se sabe que antes de la campaña, el número de accidentes diarios de transeuntes en la vía pública de la ciudad que requieren atención médica, era una variable aleatoria con distribución Poisson con esperanza igual 10. Con el fin de evaluar si la reparación de veredas ha disminuido la tasa de accidentes, durante cada uno de los 60 dias hábiles posteriores a la finalizacion de la campana de reparación, se han registrado el número de personas accidentadas en la vía pública que requirió atención médica. Sea X_i el número de accidentados en el i-ésimo día. Asuma que $X_1, \ldots, X_{60}$ forman una muestra aleatoria con distribución Poisson de parámetro λ.

(a) Plantear las hipótesis nulas y alternativas pertinentes.

(b) Calcular el test de Wald con nivel de significación asintótico igual a 0.01.

(c) Suponga que se observa una media muestral igual a 8. Qué decisión se tomaría usando el test de Wald?

Solución

(a) Las hipótesis nula y alternativa son:

H_0**:** $\lambda = 10$

H_1**:** $\lambda < 10$

(b) Por TCL sabemos que

$$\sqrt{n}\left(\overline{X}_n - \lambda\right) \xrightarrow{D} \mathcal{N}(0, \lambda),$$

entonces

$$\frac{\sqrt{n}\left(\overline{X}_n - \lambda\right)}{\sqrt{\lambda}} \xrightarrow{D} \mathcal{N}(0, 1)\,.$$

Por LGN, $\overline{X}_n \xrightarrow{p} \lambda$, y como $f(u) = \sqrt{u}$ es una función continua, $\sqrt{\overline{X}_n}$ es un estimador consistente de $\sqrt{\lambda}$. Por Slutsky,

$$\frac{\sqrt{n}\left(\overline{X}_n - \lambda\right)}{\sqrt{\overline{X}_n}} \xrightarrow{D} \mathcal{N}(0, 1)$$

Luego, bajo H_0,

$$\frac{\sqrt{n}\left(\overline{X}_n - 10\right)}{\sqrt{\overline{X}_n}} \xrightarrow{D} \mathcal{N}(0, 1)$$

Tenemos entonces que

$$W = \frac{\sqrt{n}\left(\overline{X}_n - 10\right)}{\sqrt{\overline{X}_n}},$$

donde $P\left(W < z_{0.01}\right) \to 0.01$. El test de Wald con nivel de significación asintótico 0.01 rechaza si $W < -2.32$.

Ejercicio 6. Consideremos una muestra aleatoria $X_1, \ldots, X_n$ de una distribución exponencial de parámetro λ. Derivar el test de Wald de $H_0 : \lambda = 1$ vs $H_1 : \lambda \neq 1$ con nivel de significación asintótico α.

Solución Por TCL,

$$\sqrt{n}(\overline{X}_n - \frac{1}{\lambda}) \xrightarrow{D} \mathcal{N}(0, \frac{1}{\lambda^2})\,.$$

Luego, siendo $g(u) = \frac{1}{u}$ una función continua, con $g'(u) = -\frac{1}{u^2}$, aplicando el método Delta obtenemos que

$$\sqrt{n}(\frac{1}{\overline{X}_n} - \lambda) \xrightarrow{D} \mathcal{N}(0, \lambda^2)$$

Entonces,

$$\frac{\sqrt{n}\left(\frac{1}{\overline{X}_n} - \lambda\right)}{\lambda} \xrightarrow{D} \mathcal{N}(0, 1)$$

Luego, por LGN se tiene que $\overline{X}_n \xrightarrow{p} \frac{1}{\lambda}$, con lo cual aplicando Slutsky

$$\frac{\sqrt{n}(\frac{1}{\overline{X}_n} - \lambda)}{\frac{1}{\overline{X}_n}} \xrightarrow{D} \mathcal{N}(0, 1)$$

Finalmente llamando

$$W = \frac{\sqrt{n}(\frac{1}{\overline{X}_n} - 1)}{\frac{1}{\overline{X}_n}},$$

el test rechaza H_0 con un nivel de significación α si

$$\left| \frac{\sqrt{n}\left(\frac{1}{\overline{X}_n} - 1\right)}{\frac{1}{\overline{X}_n}} \right| > z_{1-\frac{\alpha}{2}}$$

Ejercicio 7. Una cadena muy grande de comidas fast food desea investigar si la media del volumen ventas diarias en sus sucursales urbanas y sus sucursales de la periferia es la misma.

(a) Suponga que se ha tomado una muestra de sucursales urbanas de tamaño $n = 100$ y que se ha obtenido una media muestral del volumen de ventas diarias igual a \$4000 con una desviacion estandar de \$400. Suponga tambien que se ha tomado otra muestra, independiente de la primera, de sucursales urbanas de tamaño $m = 50$ y que se ha obtenido una media muestral del volumen de ventas diarias igual a \$5000 con una desviacion estandar de \$500. Conduza un test de la hipétesis nula de que las medias del volumen de ventas urbano y rural son iguales asumiendo que las varianzas de las ventas rurales y urbanas son iguales. Indique el p-valor del mismo. Si se deseara conducir un test con nivel asintótico $\alpha = 0.03$, cuál seria su conclusión?

(b) Suponga ahora que nos interesa establecer si el volumen de ventas diarias de las sucursales urbanas supera los \$4900. Nuestro interés es conducir un test de nivel $\alpha = 0.05$ de la hipótesis nula de que la media de ventas NO supera los \$4900. Asuma que el desvío estandard de la distribución del volumen de ventas diarias urbanas es menor o igual que 600. Cuál debería ser el tamaño de la muestra para que la potencia del test cuando la media verdadera es \$5000 sea aproximadamente 0.8?

Solución Empezamos resolviendo la primer parte. Las hipótesis son

$$\begin{cases} H_0 & \mu_X = \mu_Y \\ H_1 & \mu_X \neq \mu_Y \end{cases}$$

Tenemos además que $n = 100$, $\overline{x}_{100} = 4000$, $s_x = 400$, y donde $m = 50$, $\overline{y}_{50} = 5000$, $s_x = 500$.

Para estimar la varianza en común de las dos muestras, usaremos

$$S_p^2 = \frac{(n-1)S_x^2 + (m-1)S_y^2}{n+m-2},$$

por lo tanto $s_p^2 = \dfrac{99 \cdot 400^2 + 49 \cdot 500^2}{148} \approx 189797$. Cualquier otra combinación convexa de S_x^2 y S_y^2 hubiera sido válida, por ejemplo se podría haber usado $1/2S_x^2 + 1/2S_y^2$. Usamos S_p^2 porque es más eficiente.

El estadístico que usaremos es

$$\frac{\overline{X}_n - \overline{Y}_m}{\sqrt{\frac{S_p^2}{n} + \frac{S_p^2}{m}}}.$$

Su distribución asintótica bajo H_0 es:

$$\frac{\overline{X}_n - \overline{Y}_m}{\sqrt{\frac{S_p^2}{n} + \frac{S_p^2}{m}}} \overset{a}{\sim} \mathcal{N}(0,1).$$

Tenemos que que $\sqrt{\frac{s_p^2}{n} + \frac{s_p^2}{m}} \approx 75.46$ y por lo tanto el valor del estadístico observado es $\frac{1000}{75.46} \approx 13.25$.

Por lo tanto p-valor es 1-$P(-13.25 \leq Z \leq 13.25) \approx 0$. Si $\alpha = 0.03$, se rechazaría H_0.

Para la segunda parte, tenemos las hipótesis

$$\begin{cases} H_0 : & \mu \leq 4900 \\ H_1 : & \mu > 4900. \end{cases}$$

Proponemos un test con región de rechazo

$$\overline{X}_n \geq 4900 + 1.64\frac{600}{\sqrt{n}}.$$

La región de rechazo es la usual para este tipo de problemas pero 'calibrada' en el 'peor' caso, en el que $\sigma = 600$.

Veamos que el test tiene nivel asintótico 5%. Tenemos que ver que si $\mu_y \leq 4900$, cualquiera sea el σ verdadero menor o igual que 600, la probabilidad de rechazar H_0 converge a 5%. Tomemos entonces $\mu_y \leq 4900$ y $\sigma \leq 600$. La probabilidad de rechazar es

$$P_{(\mu_y,\sigma)}\left(\overline{X}_n \geq 4900 + 1.64\frac{600}{\sqrt{n}}\right) = P_{(\mu_y,\sigma)}\left(\sqrt{n}\frac{\overline{X}_n - \mu_y}{\sigma} \geq \sqrt{n}\frac{4900 - \mu_y}{\sigma} + 1.64\frac{600}{\sigma}\right).$$

Por el TCL,

$$\sqrt{n}\frac{\overline{X}_n - \mu_y}{\sigma} \overset{a}{\sim} \mathcal{N}(0,1).$$

Luego, si $Z \sim \mathcal{N}(0,1)$,

$$P_{(\mu_y,\sigma)}\left(\sqrt{n}\frac{\overline{X}_n - \mu_y}{\sigma} \geq \sqrt{n}\frac{4900 - \mu_y}{\sigma} + 1.64\frac{600}{\sigma}\right) \approx P_{(\mu_y,\sigma)}\left(Z \geq \sqrt{n}\frac{4900 - \mu_y}{\sigma} + 1.64\frac{600}{\sigma}\right).$$

Por otro lado,

$$\begin{aligned} P_{(\mu_y,\sigma)}\left(Z \geq \sqrt{n}\frac{4900 - \mu_y}{\sigma} + 1.64\frac{600}{\sigma}\right) &= 1 - P_{(\mu_y,\sigma)}\left(Z \leq \sqrt{n}\frac{4900 - \mu_y}{\sigma} + 1.64\frac{600}{\sigma}\right) \\ &= 1 - \Phi\left(\sqrt{n}\frac{4900 - \mu_y}{\sigma} + 1.64\frac{600}{\sigma}\right). \end{aligned}$$

Ahora, cómo Φ es creciente, el extremo derecho de la última igualdad es creciente en μ_y. Luego

$$\begin{aligned}1-\Phi\left(\sqrt{n}\frac{4900-\mu_y}{\sigma}+1.64\frac{600}{\sigma}\right) &\leq 1-\Phi\left(\sqrt{n}\frac{4900-4900}{\sigma}+1.64\frac{600}{\sigma}\right)\\ &= 1-\Phi\left(1.64\frac{600}{\sigma}\right).\end{aligned}$$

Nuevamente, como Φ es creciente, el extremo derecho de la última igual es creciente en σ. Luego

$$1-\Phi\left(1.64\frac{600}{\sigma}\right) \leq 1-\Phi\left(1.64\frac{600}{600}\right) = 1-\Phi(1.64) = 0.05.$$

concluimos que el test tiene nivel aproximado 5%.

Para que la potencia aproximada en 5000 sea 0.8, necesitamos un tamaño de muestra de manera que, para todo $\sigma \leq 600$

$$\begin{aligned}0.8 =&P_{(5000,\sigma)}\left(\overline{X}_n \geq 4900+1.64\frac{600}{\sqrt{n}}\right)\\ =&P_{(5000,\sigma)}\left(\overline{X}_n - 5000 \geq -100+\frac{984}{\sqrt{n}}\right)\\ =&P_{(5000,\sigma)}\left(\sqrt{n}\frac{(\overline{X}_n-5000)}{\sigma} \geq \frac{-100\sqrt{n}}{\sigma}+\frac{984}{\sigma}\right)\\ \approx&P_{(5000,\sigma)}\left(Z \geq \frac{-100\sqrt{n}}{\sigma}+\frac{984}{\sigma}\right)\\ =&1-P_{(5000,\sigma)}\left(Z \leq \frac{-100\sqrt{n}}{\sigma}+\frac{984}{\sigma}\right)\end{aligned}$$

donde en la anteúltima aproximación usamos el TCL.

Entonces

$$\frac{-100\sqrt{n}}{\sigma}+\frac{984}{\sigma} = z_{0.2} = -0.84.$$

Despejando, obtenemos que

$$\sqrt{n} = \left(\frac{990+0.84\sigma}{100}\right)$$

Como $\sigma \leq 600$, vemos que

$$\left(\frac{990+0.84\sigma}{100}\right) \leq \left(\frac{990+504}{100}\right) = 14.9$$

Luego, una muestra de tamaño $15^2 = 225$ es suficiente.

Ejercicio 8. Una muestra de 300 residentes urbanos adultos de una provincia dejó ver que 63 estaban a favor de aumentar el límite de velocidad en las rutas, mientras que una muestra, independiente de la primera, de 180 residentes rurales indicó que 75 estaban a favor del aumento. ¿Indica esta información que la opinión respecto al aumento del límite de velocidad es diferente para los dos grupos de residentes? Plantee las hipotesis nulas y alternativas adecuadas para evaluar esta pregunta y conduzca un test de las mismas con nivel asintótico $\alpha = 0.02$.

Solución

$$\begin{cases} H_0 & p_X = p_Y \\ H_1 & p_X \neq p_Y \end{cases}$$

donde $n = 300$, $\widehat{p_x} = \frac{63}{300} = \frac{21}{100}$ y donde $m = 180$, $\widehat{p_y} = \frac{75}{180} = \frac{5}{12}$.

El estadístico que utilizaremos para evaluar este test es el siguiente y que su distribución asintótica bajo H_0 es:

$$\frac{\widehat{p_x} - \widehat{p_y}}{\sqrt{\frac{\widehat{p_x}(1-\widehat{p_x})}{n} + \frac{\widehat{p_y}(1-\widehat{p_y})}{m}}} \overset{a}{\sim} \mathcal{N}(0,1)$$