Capítulo 4

INTERVALOS DE CONFIANZA DE NEYMANN

Vimos en el capítulo anterior métodos de estimación puntual. Pero no podemos esperar que la estimación que producen coincide exactamente con el verdadero valor del parámetro θ. En este capítulo buscaremos construir intervalos que proporcionen una precisión de las estimaciones encontradas. La precisión esta dada por la diferencia δ entre el estimador $\hat{\theta}$ y el parámetro θ. Si bien no conocemos a δ podemos dar como vimos anteriormente su media (el sesgo) y su varianza (el error cuadrático medio) o bien un cota u tal que es poco probable que el error $|\delta|$ sobrepase a u:

$$\mathbb{P}(|\delta| \geq u) = \mathbb{P}(|\hat{\theta} - \theta| \geq u) = \alpha$$

donde α es una probabilidad pequeña fijada a priori y $1 - \alpha$ se llama *nivel de confianza.*

Se puede escribir también $\mathbb{P}(\hat{\theta} - u \leq \theta \leq \hat{\theta} + u) = 1 - \alpha$. Si α es pequeño, se puede decir que es altamente probable de encontrar el parámetro θ en el el intervalo $[\hat{\theta} - u, \hat{\theta} + u]$, dicho de otra manera, **el intervalo $[\hat{\theta} - u, \hat{\theta} + u]$ cubre el parámetro θ con alta probabilidad $(1 - \alpha)$**.

Se dice entonces que $[\hat{\theta} - u, \hat{\theta} + u]$ es un intervalo de confianza para θ de nivel de confianza igual a $1 - \alpha$.

Observemos que aquí el intervalo $[\hat{\theta} - u, \hat{\theta} + u]$ es aleatorio y la función de distribución (ó de probabilidad) del error δ depende de la función de distribución (ó probabilidad) del estimador $\hat{\theta}$.

Apliquemos lo anterior a algunos casos.

4.1. INTERVALO PARA UNA MEDIA

Dada la función de distribución de población de la v.a. $X \sim \mathcal{N}(\theta, \sigma^2)$ y $x_1, x_2, ..., x_n$ los valores obtenidos sobre una muestra aleatoria simple de esta población, entonces la media muestral $\bar{x}_n \sim \mathcal{N}(\theta, \frac{\sigma^2}{n})$. Para construir el intervalo de confianza tenemos que considerar si la varianza σ^2 es conocida o no.

4.1.1. Caso de la varianza poblacional conocida

Si se supone que $X \sim \mathcal{N}(\theta, \sigma^2)$, se conoce la distribución del error del estimador $\bar{x}_n$ de θ: $\delta = \bar{x} - \theta \sim \mathcal{N}(0, \frac{\sigma^2}{n})$.

Para obtener el valor u dado un nivel de confianza $1-\alpha$, observamos que $\frac{\sqrt{n}}{\sigma}\delta = \frac{\sqrt{n}}{\sigma}(\bar{x}-\theta) \sim \mathcal{N}(0,1)$.

Luego usando la tabla de la distribución normal $\mathcal{N}(0,1)$ (o un programa computacional) obtenemos el valor a tal que:

$$I\!P(-a \leq Z \leq a) = 1 - \alpha$$

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

$$I\!P(\bar{x} - a\frac{\sigma}{\sqrt{n}} \leq \theta \leq \bar{x} + a\frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

El intervalo $[\bar{x} - a\frac{\sigma}{\sqrt{n}}, \bar{x} + a\frac{\sigma}{\sqrt{n}}]$ define un intervalo de confianza para la media θ de nivel de confianza $(1-\alpha)$.

Queremos que el largo del intervalo de confianza sea pequeño porque es un indicador de la precisión de la estimación, pero acá el nivel de confianza también lo es. Aquí el largo del intervalo es igual a $2a\frac{\sigma}{\sqrt{n}}$. Para n y α dados, mientras más grande es la variabilidad en la población, más largo será el intervalo. Para tener un intervalo más pequeño, es decir, con más precisión sin cambiar el nivel de confianza y la varianza σ^2, se tendrá que aumentar el tamaño n de la muestra. Si se quiere aumentar la precisión sin cambiar el tamaño de la muestra se tiene que aumentar el nivel de confianza, es decir disminuir α y tener un intervalo más grande.

Por ejemplo, para $\alpha = 0{,}05$, se obtiene el intervalo $[\bar{X} - 1{,}96\sigma/\sqrt{n}, \bar{X} + 1{,}96\sigma/\sqrt{n}]$. Pero con $\alpha = 0{,}01$, el intervalo es $[\bar{X} - 2{,}56\sigma/\sqrt{n}, \bar{X} + 2{,}56\sigma/\sqrt{n}]$. El largo del intervalo aumenta.

4.1.2. Caso de la varianza poblacional desconocida

Si no se supone que la varianza σ^2 es conocida, se tiene que usar un estadístico cuya distribución muestral no dependa de σ^2. Eso nos lleva a usar un estadístico t de Student definida en **??**. En efecto $Z = \frac{\sqrt{n}(\bar{x}-\theta)}{\sigma} \sim \mathcal{N}(0,1)$ y $U = \frac{n}{\sigma^2}s_n^2 = \frac{\sum(x_i - \bar{x})}{\sigma^2} \sim \chi^2_{n-1}$ y son independientes, luego $\frac{Z}{\sqrt{u/(n-1)}} \sim t_{n-1}$. El estadístico depende del error δ, por lo tanto de θ, pero no de la varianza σ^2 desconocida:

$$T = \frac{\bar{x} - \theta}{s_n/\sqrt{(n-1)}}$$

que sigue una distribución t Student a $n-1$ grados de libertad.

Procedemos como en el caso de la varianza conocida. Buscamos el valor a para una probabilidad pequeña α dado tal que

$$I\!P(-a \leq t_{n-1} \leq a) = 1 - \alpha$$

Se deduce el intervalo

$$[\bar{x} - a\frac{s_n}{\sqrt{n-1}}, \bar{x} + a\frac{s_n}{\sqrt{n-1}}]$$

para θ de nivel de confianza $1-\alpha$.

4.2. INTERVALO PARA LA VARIANZA

Si los valores muestrales $x_1, x_2, ..., x_n$ son i.i.d. de la $\mathcal{N}(\theta, \sigma^2)$, el estimador de máxima verosimilitud de σ^2 es $\hat{\sigma}^2 = \frac{1}{n}\sum(x_i-\theta)^2$ si la media θ es conocida y $s_n^2 = \frac{1}{n}\sum(x_i-\bar{x})^2$ si θ es desconocida. El estadístico que se utilizará en la construcción del intervalo de confianza depende entonces de si la media poblacional es conocida o no.

4.2.1. Caso de la media poblacional conocida

En este caso se usa el estadístico $U_n = n\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2_n$. Considerando que si

$$I\!P(u_1 \leq U_n \leq u_2) = 1-\alpha$$

entonces

$$I\!P(n\frac{\hat{\sigma}^2}{u_2} \leq \sigma^2 \leq \frac{\hat{\sigma}^2}{u_1}) = 1-\alpha$$

expresión que nos permite construir un intervalo para σ^2 de nivel de confianza igual a $1-\alpha$.

El problema aquí es que nos sabemos como elegir el par de cotas (u_1,u_2) que nos son únicas con el mismo nivel de confianza $1-\alpha$. Considerando que el largo del intervalo para un nivel de confianza $1-\alpha$ dado esta relacionado con la precisión de la estimación, lo mejor sería encontrar el intervalo más pequeño entre todos los intervalos de mismo nivel de confianza. Pero este problema no tiene una solución simple. En la práctica lo que lleva a simplificar el problema es tomar las cotas u_1 y u_2 tales que:

$$I\!P(U_n \leq u_1) = I\!P(U_n \geq u_2) = \frac{\alpha}{2}$$

Por ejemplo, con una muestra de tamaño $n=20$ para $\alpha = 0{,}05$, $U_1 = 9{,}59$ y $U_2 = 34{,}17$, se obtiene el intervalo $[20\frac{\hat{\sigma}^2}{34{,}17}, 20\frac{\hat{\sigma}^2}{9{,}59}]$. Pero con $\alpha = 0{,}01$, el intervalo es $[20\frac{\hat{\sigma}^2}{39{,}90}, 20\frac{\hat{\sigma}^2}{7{,}43}]$.

El largo del intervalo aumenta cuando se disminuye α y disminuye cuando se aumenta n.

4.2.2. Caso de la media poblacional desconocida

Ahora se usa el estadístico $U_{n-1} = \frac{\sum(X_i-\bar{X})^2}{\sigma^2} \sim \chi^2_{n-1}$.

Considerando que si

$$I\!P(u_1 \leq U_{n-1} \leq u_2) = 1-\alpha$$

entonces

$$I\!P(n\frac{s_n^2}{u_2} \leq \sigma^2 \leq \frac{s_n^2}{u_1}) = 1 - \alpha$$

expresión que nos permite construir un intervalo para σ^2 de nivel de confianza igual a $1-\alpha$ cuando la media poblacional es desconocida. Como en el caso anterior se elige u_1 y u_2 de manera que:

$$I\!P(U_{n-1} \leq u_1) = I\!P(U_{n-1} \geq u_2) = \frac{\alpha}{2}$$

o sea

$$I\!P(\frac{\sum(x_i - \bar{x})^2}{u_2} \leq \sigma^2 \leq \frac{\sum(x_i - \bar{x})^2}{u_1}) = 1 - \alpha$$

Por ejemplo, con una muestra de tamaño $n = 20$ para $\alpha = 0{,}05$, $U_1 = 8{,}91$ y $U_2 = 32{,}85$, se obtiene el intervalo $[20\frac{\hat{\sigma}^2}{32{,}85}, 20\frac{\hat{\sigma}^2}{8{,}91}]$. Pero con $\alpha = 0{,}01$, el intervalo es $[20\frac{\hat{\sigma}^2}{38{,}58}, 20\frac{\hat{\sigma}^2}{8{,}84}]$.

4.3. LA DIFERENCIA DE DOS MEDIAS

Sean dos poblaciones normales $\mathcal{N}(\mu_1, \sigma_1^2)$ y $\mathcal{N}(\mu_2, \sigma_2^2)$. Se considera una muestra aleatoria simple de tamaño n_1 de la primera población y una muestra aleatoria simple de tamaño n_2 de la segunda población, las dos muestras siendo independientes. Si $\bar{x}_1$ y $\bar{x}_2$ son las medias muestrales respectivas, $d = \bar{x}_1 - \bar{x}_2 \sim \mathcal{N}(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$.

Si las varianzas son conocidas entonces un intervalo para δ la diferencia de las medias de las poblaciones esta dado por:

$$[\bar{x}_1 - \bar{x}_2 - u\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, \quad \bar{x}_1 - \bar{x}_2 + u\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}]$$

con u determinado a partir de las tablas de la distribución normal según el nivel de confianza $1-\alpha$.

Si las varianzas no son conocidas, para encontrar un estadístico que sirva y cuya distribución no dependa de estas varianzas, hay que hacer alguno supuesto suplementario. En efecto si tomamos como estimador de la varianza de la diferencia $\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}$ con s_1^2 y s_2^2 las varianzas muestrales sesgadas, $\frac{n_1 s_1^2}{\sigma_1^2} + \frac{n_2 s_2^2}{\sigma_2^2} \sim \chi^2_{n_1+n_2-2}$ y

$$\frac{(\bar{s}_1 - \bar{s}_2 - (\mu_1 - \mu_2))/\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}{\sqrt{(\frac{n_1 s_1^2}{\sigma_1^2} + \frac{n_2 s_2^2}{\sigma_2^2})/(n_1 + n_2 - 2)}} \sim t_{n_1+n_2-2}$$

que depende de la varianzas desconocidas σ_1^2 y σ_2^2.

Si se supone que estas varianzas son proporcionales: $\sigma_2^2 = k^2\sigma_1^2$, entonces se tiene un estadístico que no depende de σ_1^2 y σ_2^2:

$$\frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{(\frac{k^2 n_1 s_1^2 + n_2 s_2^2}{k^2(n_1+n_2-2)})(\frac{k^2 n_1 + n_2}{n_1 n_2})}} \sim t_{n_1+n_2-2}$$

Usualmente si toma $k = 1$. Se obtiene el valor de u tal que

$$I\!P(-u \leq t_{n_1+n_2-2} \leq u) = 1 - \alpha$$

En este caso el intervalo para $\mu_1 - \mu_2$ es:

$$[\bar{x}_1 - \bar{x}_2 - u\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}\sqrt{\frac{n_1 + n_2}{n_1 n_2}}, \quad \bar{x}_1 - \bar{x}_2 + u\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}\sqrt{\frac{n_1 + n_2}{n_1 n_2}}$$

Sea, por ejemplo, $n_1 = 15$, $n_2 = 17$, $\bar{x}_1 = 4$, $\bar{x}_2 = 5$, $S_1^2 = 2{,}0$, $S_2^2 = 2{,}5$ y $\alpha = 0{,}05$. En este caso $u = 2{,}042$ y el intervalo es

$$[-1 - 2{,}042\sqrt{\frac{72{,}5}{30}}\sqrt{\frac{32}{255}}, \quad -1 + 2{,}042\sqrt{\frac{72{,}5}{30}}\sqrt{\frac{32}{255}} = [-2{,}1247, \quad 0{,}1247]$$

Se observa que el intervalo cubre apenas el valor 0, lo que indica que las medias son bastante distintas.

4.4. EL COCIENTE DE DOS VARIANZAS

En el párrafo anterior para construir el estadístico t de Student se hizo el supuesto que las varianzas σ_1^2 y σ_2^2 son proporcionales. Este supuesto podra verificarse usando un intervalo de confianza para el cociente de las dos varianzas $\frac{\sigma_1^2}{\sigma_2^2}$.

Sean dos poblaciones normales $\mathcal{N}(\mu_1, \sigma_1^2)$ y $\mathcal{N}(\mu_2, \sigma_2^2)$ y una muestra aleatoria simple de cada población, ambas tomadas de manera independiente. Nos interesamos en el cociente de las dos varianzas: $\dfrac{\sigma_1^2}{\sigma_2^2}$.

El estadístico $n_1 s_1^2/\sigma_1^2 \sim \chi^2_{n_1-1}$ y el estadístico $n_2 s_2^2/\sigma_2^2 \sim \chi^2_{n_2-1}$, siendo estos independientes. Vamos a definir una nueva distribución para el cociente de dos varianzas empíricas llamada distribución F de Fisher.

Definición 4.4.1 *El cociente de una variable χ^2 a r grados de libertad y de una variable χ^2 a s grados de libertad, independientes entre si, sigue una distribución llamada F de Fisher a r y s grados de libertad. Se denota $F_{r,s}$.*

Mostramos que $F_{r,s}$ tiene una función de densidad igual a:

$$h(y) = \frac{\Gamma(\frac{r+s}{2})}{\Gamma(\frac{r}{2})\Gamma(\frac{s}{2})} \frac{r^{r/2}s^{s/2}y^{(r/2)-1}}{(ry+s)^{(r+s)/2}} \quad \forall y > 0$$

Sean $U \sim \chi_r^2$ y $V \sim \chi_s^2$ independientes entre si. Como U y V son independientes, se puede calcular fácilmente la función de densidad conjunta de (U,V):

$$f(u,v) = \frac{u^{(r/2)-1}e^{-u/2}}{2^{r/2}\Gamma(r/2)} \frac{v^{(s/2)-1}e^{-v/2}}{2^{s/2}\Gamma(s/2)}$$

Con el cambio de variables:$(U,V) \longrightarrow (Y,Z)$ con $U = rYZ/s$ y $V = Z$, obtenemos la densidad conjunta de (Y,Z):

$$g(y,z) = \frac{(r/s)z}{2^{(r+s)/2}\Gamma(r/2)\Gamma(s/2)}(r/s)^{(r/2)-1}y^{(r/2)-1}z^{(r+s-1)/2}e^{-1/2(ry/s+1)z}$$

Se deduce la densidad marginal de Y:

$$f(y) = \int_0^\infty g(y,z)dz = \frac{\Gamma(\frac{r+s}{2}r^{r/2}s^{s/2}y^{(r/2)-1}}{\Gamma(r/2)\Gamma(s/2)(ry+s)^{(r+s)/2}}$$

Nota 4.4.2 *Observamos que si* $Y \sim F_{r,s}$ *entonces* $1/Y \sim F_{s,r}$*. Además* $\dfrac{rY/s}{1+rW/s} \sim$ $\beta eta((r-2)/2,(s-2)/2)$*.*

Aquí el estadístico $\dfrac{n_1S_1^2/(n_1-1)\sigma_1^2}{n_2S_2^2/(n_2-1)\sigma_2^2} \sim F_{n_1-1,n_2-1}$, lo que permite construir un intervalo de confianza para el cociente σ_1^2/σ_2^2.

Si $I\!P(a \leq F_{n_1-1n2-1} \leq b) = 1-\alpha$, entonces el intervalo es

$$[\frac{n_1s_1^2(n_2-1)}{bn_2s_2^2(n_1-1)}, \quad \frac{n_1s_1^2(n_2-1)}{an_2s_2^2(n_1-1)}]$$

4.5. INTERVALO PARA LA PROPORCIÓN

Sea la proporción θ de piezas defectuosas en un lote de piezas fabricadas por la industria **??**. El número de piezas defectuosas encontradas en una muestra aleatoria simple de tamaño n sigue una distribución binomial $B(n,\theta)$. Construir un intervalo de confianza para una proporción es más complicado que construirlo para una media o varianza. Cuando n es pequeño hay que usar la distribución binomial (tablas y ábacos fueron calculados para determinar valores de θ_1 y θ_2 para los diferentes valores de k y n y de nivel de confianza $1-\alpha$).

Cuando n es grande, se puede usar la aproximación a la distribución normal $\mathcal{N}(n\theta, n\theta(1-\theta))$, pero subsiste un problema ya que la varianza depende también de θ. Una manera de proceder considerando la proporción empírica $\hat{p} \approx \mathcal{N}(\theta, \frac{\theta(1-\theta)}{n})$. Se tiene entonces:

$$I\!P(\mid \frac{\sqrt{n}(\hat{p}-\theta)}{\sqrt{\theta(1-\theta)}} \mid \leq u) = 1-\alpha$$

Lo que equivale a:

$$I\!P(n(\hat{p}-\theta)^2 - u^2\theta(1-\theta) \leq 0) = 1-\alpha$$

Las soluciones de la ecuación:

$$(n+u^2)\theta^2 - (2n\hat{p}+u^2)\theta + n\hat{p}^2 = 0$$

siendo $\dfrac{2n\hat{p}+u^2 \pm \sqrt{u^4+4n\hat{p}u^2-4nu^2\hat{p}^2}}{2(n+u^2)}$, se obtiene el intervalo:

$$[\frac{n}{n+u^2}(\hat{p}+\frac{u^2}{2n}) - u\sqrt{\frac{\hat{p}(1-\hat{p})}{n}+\frac{u^2}{4n^2}}, \quad \frac{n}{n+u^2}(\hat{p}+\frac{u^2}{2n}) + u\sqrt{\frac{\hat{p}(1-\hat{p})}{n}+\frac{u^2}{4n^2}}]$$

Para n grande, se puede aproximar por:

$$[\hat{p} - u\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \quad \hat{p} + u\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}]$$

4.6. EJERCICIOS

1. Sea una m.a.s. $\{x_1, ...x_n\}$ de una distribución normal de media θ desconocida y varianza σ^2 conocida.

a) Dé el número mínimo n del tamaño de la muestra para que un intervalo de confianza I a 95 % tenga un largo L a lo más igual a 0.016 σ.
b) Sea $L = \sigma/5$. Dé el nivel de confianza $1-\alpha$ cuando n=10, 20, 30 y 100.
c) Repetir b) con σ^2 desconocido. Comente.
d) Dé el intervalo de confianza de largo mínimo para θ con un nivel de confianza de 95 %, cuando $\sigma^2 = 4$.

2. Una empresa desea estimar el promedio de tiempo que necesita una secretaria para llegar a su trabajo. Se toma una m.a.s. de 36 secretarias y se encuentra un promedio de 40 minutos. Suponiendo que el tiempo de trayecto proviene de una $\mathcal{N}(\mu, \sigma^2)$, con $\sigma = 12$, dé un intervalo de confianza para la media μ.

3. Se dispone de 10 muestras de sangre tomadas en las mismas condiciones a una misma persona. Se obtiene para cada una la dosis de Colesterol (en gramos) *245, 248, 250, 247, 249, 247, 247, 246, 246, 248*. Cada medida puede considerarse como una realización particular de la variable "tasa de Colesterol" $X \sim \mathcal{N}(\mu, \sigma^2)$.

a) Dé un intervalo de confianza para μ al 95 % suponiendo $\sigma^2 = 1{,}5$.
b) Dé un intervalo de confianza para μ al 95 % suponiendo σ^2 desconocido.
c) Construya un intervalo de confianza para σ^2 al 95 % .

4. Dé una inecuacón para determinar el intervalo de confianza del parámetro θ de una variable de Bernoulli.

5. Se tienen 2 muestras de tamaños n_1 y n_2 de una misma v.a. X medida sobre dos poblaciones distintas. Se asume que para ambas poblaciones X sigue una distribución Normal con medias μ_1, μ_2 y varianzas σ_1^2, σ_2^2, respectivamente.

a) Construya un intervalo de confianza para $\mu_1 - \mu_2$, suponiendo que $\sigma_2^2 = k^2\sigma_1^2$ en que k es una constante conocida.
b) Muestre que los extremos del intervalo anterior convergen en probabilidad si los tamaños de las muestras crecen.
c) Se supone ahora la constante k desconocida. Dé un método para construir un intervalo de confianza para la constante k.
d) ¿ Que inconveniente cree Ud. que tiene este método?

6. Considere una v.a. $X \sim \mathcal{N}(\mu, 1)$ y una m.a.s. de X con una sóla observación x. Dada una constante $a > 0$, se define el intervalo aleatorio: $C_a(x) = [min(0, x - a), max(0, x + a)]$.

a) Muestre que $I\!P(\mu \in C_a(x)/\mu = 0) = 1 \ \forall x$.
b) Muestre que $C_a(x)$ es un intervalo de confianza para μ de nivel de confianza $1 - \alpha = 95\,\%$, cuando a=1.65.
c) Sea $\pi(\mu) = 1$ $(\forall \mu)$ una distribución a priori para μ. Deducir la distribución a posteriori de μ dado x.
d) Sea Φ la función de distribución de la normal $\mathcal{N}(0, 1)$. Muestre que se encuentra una probabilidad condicional

$$I\!P(\mu \in C_a(x)/x) = \begin{cases} \Phi(-x) - \Phi(-a) & \text{si } x < -a \\ \Phi(a) - \Phi(-a) & \text{si } -a < x < a \\ \Phi(a) - \Phi(-x) & \text{si } x > a \end{cases}$$

e) Deducir que, para a=1.65, la probabilidad condicional $I\!P(\mu \in C_a(x)/x) \geq 0{,}90$ y que $\lim_{a\to\infty} I\!P(\mu \in C_a(x)/x) = 1$.

Capítulo 5

TESTS DE HIPOTESIS

5.1. ¿COMÓ UN JUEZ SENTENCIA?

Sir Ronald Fisher (párrafo **??**) describe en su escrito Experimental Design (1935) lo siguiente: Una dama británica (Lady, en el texto) declara ser capaz de distinguir si la leche fue puesta antes o después del té en su taza. Fisher propone entonces un experimento para comprobar lo que dice esta dama. Se prepara 4 tazas de té en las cuales se puso la leche antes del té y 4 otras tazas en las cuales se puso la leche después del té. Se presentan al azar las 8 tazas a la Dama, que las prueba todas y decide cuáles tuvieron la leche antes y cuáles tuvieron la leche después. Si acierta a las 8 tazas, ¿diría Ud que fue solamente suerte? ¡Nosotros, diríamos que no! En efecto tenemos repuestas posibles y una sola repuesta correcta. Por lo tanto, si la dama ha contestado al azar, tenía una probabilidad de 1/70 de dar la repuesta correcta. ¿Es eso suerte?

En el capítulo 3, se presentaron métodos que permiten encontrar los valores de los parámetros desconocidos de la distribución de población y en el capítulo anterior, la estimación por intervalo permite dar una cierta indicación sobre la **precisión** de la estimación puntual. Tales estimaciones, puntuales y por intervalo, que fueron obtenidas a partir de valores muestrales, permiten formarse una opinión sobre la población y entonces darse una **hipótesis** de trabajo.

Ejemplos:

- Antes de apostar "cara" o "sello" en el lanzamiento de una moneda, se tiene que postular que la moneda está equilibrada. La hipótesis de trabajo es entonces que el parámetro p=probabilidad de sacar "cara" de la distribución de Bernoulli es

$$p = 0{,}5$$

- Un agricultor se compromete a entregar a una fábrica de azúcar remolacha con un cierto porcentaje p_o de glucosa; la hipótesis de trabajo es entonces

$$p = p_o \qquad o \qquad p \geq p_o$$

- Los hombres chilenos pretenden ser más altos que los argentinos en promedio; si μ_1 y μ_2 son las tallas promedias respectivas de los hombres chilenos y argentinos, la hipótesis de trabajo es
$$\mu_1 \geq \mu_2$$
- En el ejemplo sobre la estimación puntual de la talla promedia μ_1 de los hombres chilenos, se hizo la hipótesis de trabajo que la v.a. X talla de los hombres chilenos sigue una distribución
$$F \sim Normal$$

En los cuatro casos se procederá de la misma manera: se tiene una hipótesis de trabajo y un experimento que nos proporciona una muestra de observaciones; se trata de decidir si la hipótesis planteada es compatible con lo que se puede aprender del estudio de los valores muestrales. Se tiene que encontrar un procedimiento para decidir si la muestra que se obtuvo esta en acuerdo con la hipótesis de trabajo. Naturalmente no se espera que, para cualquier muestra, el valor empírico obtenido en la muestra coincide con el valor esperado de la hipótesis; el problema es entonces decidir si la desviación encontrada entre el valor esperado y el valor observado en la muestra es demasiado grande para poner en duda la hipótesis de trabajo. Ahora bien si se pone en duda la hipótesis original, entonces se la rechaza en favor de una **hipótesis alternativa**.

En efecto, en el ejemplo de la moneda, si se encuentra en 100 lanzamientos un 45 % de caras, ¿debemos rechazar la hipótesis $p = 1/2$,la moneda esta equilibrada? y si se rechaza esta hipótesis, ¿será a favor de la hipótesis $p \leq 1/2$, la moneda esta cargada en los "sellos"?

Se distingue la hipótesis de trabajo llamándola **hipótesis nula** y la cual se confronta a una **hipótesis alternativa**.

> *¿Con qué grado de desacuerdo uno tiene que abandonar la hipótesis nula a favor de la hipótesis alternativa?*

Para decidir entre las dos hipótesis, se necesita una **regla de decisión**.

Cualquiera regla de decisión debería tratar de minimizar los errores de decisión. Si δ es la regla de decisión adoptada y $\alpha(\delta)$ la probabilidad de equivocarse cuando la hipótesis nula es cierta y $\beta(\delta)$ la probabilidad de equivocarse cuando la hipótesis alternativa es cierta, uno buscara minimizar ambas probabilidades de equivocarse. Pero veremos, a través de un ejemplo, que no se puede disminuir los dos errores al mismo tiempo, en particular para tener $\alpha(\delta)$ nula se hace $\beta(\delta)$ igual a 1 viceversa.

Entonces para decidir entre la hipótesis nula y la hipótesis alternativa se corre el riesgo de obtener resultados falsos y tomar decisiones incorrectas con consecuencias graves. ¿Qué riesgo estamos dispuestos a asumir?. ¿Los dos errores tienen la misma importancia para la decisión a tomar? Se puede entender un poco mejor el procedimiento de un test de hipótesis haciendo el paralelo con las prácticas procesales en el sistema judicial. Frente a un acusado, el juez tiene que tomar la decisión grave entre dos casos: "El acusado es inocente" o "el acusado es culpable". El juez escucha el abogado de la acusación, que recopiló los antecedentes del

caso en contra del acusado y escucha la defensa que trata de refutar las evidencias de la acusación.

La acusación debe presentar suficiente evidencia que permitan convencer al juez de la culpabilidad del acusado, al menos más allá de “una duda razonable”. El resultado por defecto, si no hay suficiente pruebas de culpabilidad, es que ”el acusado es inocente”. Solo en el caso de una evidencia abrumadora, el juez declarará ”el acusado culpable”. El veredicto será automáticamente ”inocente”. Es el rol de la acusación de tener un caso para culpabilizar el acusado y no a la defensa de probar la inocencia, pero refutar la evidencia presentada por la acusación.

Ahora si el juez declara culpable a un inocente, no prueba que el acusado sea realmente culpable. Pueden ocurrir errores por una mala defensa. Sin embargo, los procedimientos de investigación y de justicia son establecidos para controlar la probabilidad que tal error ocurra. Por otra parte, si un acusado es absuelto, no significa que es inocente. Significa que la acusación no encontró evidencia adecuada de su culpabilidad.

El juez puede cometer dos tipos de errores, pero no se le da la misma importancia a ambas. El juez va a tratar de minimizar el error de *condenar a un inocente*.

Dada una hipótesis nula H_o, vimos que $\alpha(\delta)$ es la probabilidad de rechazar la hipótesis H_o con la regla δ cuando H_o es cierta. Ahora bien la regla δ se basa en los valores muestrales, es decir en evidencias; si la muestra es de tamaño n y los valores muestrales pertenecientes a Q, una regla de decisión δ consiste en dividir el dominio Q^n del conjunto de todas las muestras de tamaño n en dos partes disjuntas: la parte W en donde las muestras conducen a rechazar la hipótesis nula H_o, es decir evidencias contra el acusado presentadas por la acusación, y la parte $\overline{W}$ en donde no se rechaza H_o, evidencias presentadas por la defensa que recusan las evidencias de la acusación. La parte W se llama **región de rechazo de** H_o o **región crítica del test**, que consiste en las evidencias que permiten declarar culpable al acusado.

Como la región crítica del test es aquella en donde se rechaza H_o, debería tomar en cuenta la hipótesis alternativa.

Una regla de decisión consiste entonces en determinar la región crítica del test en función de las dos hipótesis.

5.2. HIPÓTESIS ESTADÍSTICAS

Las hipótesis estadísticas son muy precisas: se refieren al comportamiento de variables aleatorias. Pero en los ejemplos expuestos en el párrafo anterior, se observó que las hipótesis no son todas del mismo tipo. En los tres primeros ejemplos, la hipótesis concierne solamente a los valores de parámetros de una distribución cuya forma es especificada a priori y no está puesta en duda. Tales hipótesis se llaman **hipótesis parámetricas**. En el último ejemplo "$F \sim Normal$", es la distribución completa que esta puesta en juicio; se habla de **hipótesis no parámetricas**.

Por ejemplo, sea una v.a. X de distribución $F_\theta(x)$, que depende de un parámetro θ. Si Ω es el espacio del parámetro θ y Ω_o un subconjunto de Ω, entonces

$$H : \theta \in \Omega_o$$

es una hipótesis parámetrica, mientras que

$$H : F \sim Normal$$

es una hipótesis no parámetrica.

Se puede clasificar también las hipótesis parámetricas según su grado de especificidad. Cuando en la hipótesis parámetrica $H_o : \theta \in \Omega_o$, Ω_o esta reducido a un sólo valor, entonces se habla de **hipótesis simple**, sino se habla de **hipótesis compuesta**.

5.3. TEST DE HIPÓTESIS PARÁMETRICAS

Trataremos en primer lugar los tests de hipótesis parámetricas para hipótesis simples antes de tratar el caso general apoyandonos en los resultados del caso de las hipótesis simples. Encontrar una regla de decisión es encontrar una región crítica del test. ¿Como hacerlo minimizando los errores de decisión? Para eso usaremos la función de potencia.

5.3.1. Función de potencia

Sea un test de hipótesis sobre el parámetro θ ($\theta \in \Omega$) de la distribución F de una v.a. X.

$$H_o : \theta \in \Omega_o \quad contra \quad H_1 : \theta \in \Omega_1$$

Si una regla de decisión nos condujo a una región crítica W para el test, entonces para cada valor de $\theta \in \Omega$, determinaremos la probabilidad $\pi(\theta)$ que la regla de decisión definida por W nos conduce a rechazar H_o cuando el parámetro vale θ.

Definición 5.3.1 *La FUNCIÓN DE POTENCIA DEL TEST es*

$$\pi(\theta) = \mathbb{P}(decidir \quad rechazar \quad H_o|\theta)$$

¡OJO! aquí θ no es una variables aleatoria.

W es la región crítica del test y $\underline{x}$ el vector de los valores muestrales, entonces

$$\pi(\theta) = I\!P(\underline{x} \in W|\theta) \quad \forall\theta \in \Omega$$

Luego la región crítica ideal es aquella que produce una función de potencia tal que:

$$\pi(\theta) = \begin{cases} 0 & \text{si } \theta \in \Omega_o \\ 1 & \text{si } \theta \in \Omega_1 \end{cases}$$

En efecto, para todo $\theta \in \Omega_o$, la decisión de rechazar H_o es una decisión equivocada, entonces $\pi(\theta)$ es **una probabilidad de error de tipo I** (ó riesgo de primer especie). Por otro lado, para todo $\theta \in \Omega_1$, la decisión de rechazar H_o es una decisión correcta, entonces $1 - \pi(\theta)$ es **una probabilidad de error de tipo II** (ó riesgo de segundo especie).

Definición 5.3.2 *Se llama TAMAÑO del test a* $sup\{\pi(\theta)|\theta \in \Omega_o\}$

Veamos en un ejemplo que una región crítica ideal, con $\pi(\theta) = 0$ en Ω_o y $\pi(\theta) = 1$ en Ω_1, no existe.

Sea $x_1, x_2, ..., x_n$ una m.a.s. de una v.a. X uniforme en $[0,\theta]$ con $\theta > 0$.
Consideramos la hipótesis nula $H_o : 3 \leq \theta \leq 4$ contra la hipótesis alternativa $H_1 : \theta < 3$ o $\theta > 4$. Supongamos que una regla de decisión δ nos llevo a decidir de no rechazar a la hipótesis nula H_o cuando $\max\{x_1, x_2, ..., x_n\}$ de una m.a.s. de la v.a. X esta en el intervalo [2.9,4.1] y a rechazar H_o en el caso contrario. Luego la región crítica del test es un subconjunto $W \subset I\!R^n$ tal que $\max\{x_1, x_2, ..., x_n\} < 2{,}9$ o $> 4{,}1\}$. La función de potencia del test es entonces:

$$\boxed{\pi(\theta) = I\!P(\max\{x_1 + x_2, ..., x_n\} < 2{,}9|\theta) + I\!P(\max\{x_1, x_2, ..., x_n\} > 4{,}1|\theta)}$$

Si $\theta \leq 2{,}9 \Rightarrow \left\{ \begin{array}{l} I\!P(\max\{x_1, x_2, ..., x_n\} < 2{,}9|\theta) = 1 \\ I\!P(\max\{x_1, x_2, ..., x_n\} > 4{,}1|\theta) = 0 \end{array} \right\} \Rightarrow \pi(\theta) = 1$

Si $2{,}9 < \theta \leq 4{,}1 \Rightarrow \left\{ \begin{array}{l} I\!P(\max\{x_1, x_2, ..., x_n\} < 2{,}9|\theta) = (\frac{2{,}9}{\theta})^n \\ I\!P(\max\{x_1, x_2, ..., x_n\} > 4{,}1|\theta) = 0 \end{array} \right\} \Rightarrow \pi(\theta) = (\frac{2{,}9}{\theta})^n$

Si $\theta > 4{,}1 \Rightarrow \left\{ \begin{array}{l} I\!P(\max\{x_1, x_2, ..., x_n\} < 2{,}9|\theta) = (\frac{2{,}9}{\theta})^n \\ I\!P(\max\{x_1, x_2, ..., x_n\} > 4{,}1|\theta) = 1 - (\frac{4{,}1}{\theta})^n \end{array} \right\} \Rightarrow \pi(\theta) = 1 + (\frac{2{,}9}{\theta})^n - (\frac{4{,}1}{\theta})^n$

El tamaño del test es igual a $\alpha = Sup\{\pi(\theta)|3 \leq \theta \leq 4\} = \pi(3) = (\frac{2{,}9}{3})^n$

En los gráficos 5.1, se muestra la función de potencia para los casos $n = 10$ y $n = 50$. Se observa que el tamaño del test $\alpha = 0{,}10$, es decir que en el intervalo $[3, 4]$ la probabilidad de equivocarse no sobrepasa 10 %. Pero el error de tipo II, que es igual a $1 - \pi(\theta)$ cuando $\theta \in \Omega_o$, puede ser muy elevado; entre 3 y 2,9, el error disminuye de 10 % a 0 %; pero entre 4 y 4,1 es casi igual a 1.

En este ejemplo si queremos disminuir el tamaño del test α, hay que elegir un intervalo $\overline{W}$ más grande o una muestra de tamaño mayor. Pero en ambos casos se aumentara el error de tipo II. Para tratar de acercarnos a la situación ideal, se puede, por ejemplo, buscar minimizar una función de los dos errores, o bien fijarse una cota máxima para el error de tipo I y minimizar el error de tipo II, como en el caso del juez.

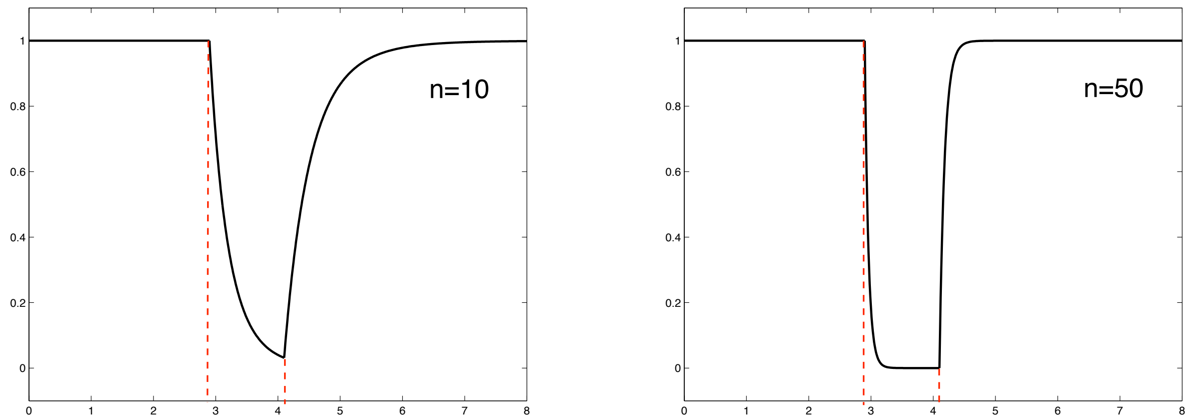

Figura 5.1: Función de potencia para la región crítica [2,9, 4,1]

5.3.2. Tests para hipótesis simples

Sean $x_1, x_2, ..., x_n$, los valores muestrales independientes de una v.a. de función de densidad $f_\theta(x)$. Se plantea las hipótesis simples:

$$H_o : \theta = \theta_o \quad contra \quad H_1 : \theta = \theta_1$$

Dada una regla de decisión δ, se tienen los dos errores:

$$\alpha(\delta) = I\!P(rechazar H_o|\theta = \theta_o) \quad (error\ de\ tipo\ I)$$

$$\beta(\delta) = I\!P(no\ rechazar H_o|\theta = \theta_1) \quad (error\ de\ tipo\ II)$$

Tenemos dos estrategias para tratar los dos errores:

(a)Minimizar una función simple de los dos errores, como el promedio ponderado de los errores tipo I y tipo II $a\alpha(\delta) + b\beta(\delta)$, con $a + b = 1$.

(b)Fijarse una cota máxima α_o para el error de tipo I y minimizar el error de tipo II.

Mostraremos en primer lugar como minimizar el promedio de los dos errores. Para el según caso usaremos la solución de (a) para encontrar la forma de construir la región crítica en el caso (b).

Dados dos escalares a y b, buscamos minimizar la función $a\alpha(\delta) + b\beta(\delta)$. Se denota $f_o(\underline{x})$ y $f_1(\underline{x})$ a las funciones de verosimilitud dado H_o y dado H_1 respectivamente:

$$f_o(\underline{x}) = \prod_i^n f(x_i/\theta_o) \quad y \quad f_1(\underline{x}) = \prod_i^n f(x_i/\theta_1)$$

Teorema 5.3.3 *Si δ^* es la regla de decisión tal que:*

$$\begin{array}{c} \textit{se rechaza } H_o \textit{ cuando } af_o(\underline{x}) < bf_1(\underline{x}), \\ \textit{se acepta } H_o \textit{ cuando } af_o(\underline{x}) > bf_1(\underline{x}), \end{array}$$

entonces $a\alpha(\delta^*) + b\beta(\delta^*) \leq a\alpha(\delta) + b\beta(\delta) \ \forall\delta$

Demostración Si W es la región crítica asociada a una regla de decisión δ,

$$\alpha(\delta) = \int ... \int_W f_o(\underline{x})dx_1...dx_n$$

$$\beta(\delta) = \int ... \int_{\overline{W}} f_1(\underline{x})dx_1...dx_n$$

$$a\alpha(\delta) + b\beta(\delta) = a\int ... \int_W f_o(\underline{x})dx_1...dx_n + b(1 - \int ... \int_W f_1(\underline{x})dx_1...dx_n)$$

Luego $a\alpha(\delta) + b\beta(\delta)$ es mínimo cuando $\int ... \int_W (af_o(\underline{x}) - bf_1(\underline{x}))dx_1...dx_n$ es mínimo.
Es decir si: $\begin{cases} af_o(\underline{x}) - bf_1(\underline{x}) < 0 \ \forall \underline{x} \in W \\ af_o(\underline{x}) - bf_1(\underline{x}) > 0 \ \forall \underline{x} \in \overline{W} \end{cases}$

entonces δ^* es óptimo para estos valores a y b dados. Se observará que $f_o(\underline{x}) - bf_1(\underline{x}) = 0$ es irrelevante, dado que no cambia el mínimo. ■

Para la segunda estrategía (b) supongamos que α_o la cota máxima de error de tipo I que se quiere aceptar. Daremos en primer lugar definiciones.

Definición 5.3.4 *Se llama RAZÓN DE VEROSIMILITUD de la muestra al cociente*

$$\frac{f_1(\underline{x})}{f_o(\underline{x})}$$

Definición 5.3.5 *Se llama NIVEL DE SIGNIFICACIÓN del test a la cota máxima de error de tipo I aceptada.*

Se tiene entonces que buscar una regla de decisión δ que produzca un error de tipo I $\alpha(\delta) \leq \alpha_o$ y tal que $\beta(\delta)$ sea mínimo. El siguiente lema, que deriva del teorema anterior, nos da la forma de proceder.

Lema 5.3.6 *(NEYMAN-PEARSON)*
Si δ^* es una regla de decisión tal que para algún $k > 0$ fijo,
se rechaza H_o, si $\dfrac{f_1(\underline{x})}{f_0(\underline{x})} > k$
no se rechaza H_o, si $\dfrac{f_1(\underline{x})}{f_0(\underline{x})} < k,$
entonces para toda regla δ tal que $\alpha(\delta) \leq \alpha(\delta^*)$ se tiene $\beta(\delta) \geq \beta(\delta^*)$.

Sea $x_1, ..., x_n$ una muestra aleatoria simple de la v.a. $X \sim \mathcal{N}(\mu, \sigma^2)$, μ desconocido y σ^2 conocido. Se estudia $H_o : \mu = 1$ contre $H_1 : \mu = 2$. La razón de verosimilitud se escribe:

$$\frac{f_1(\underline{x})}{f_o(\underline{x})} = exp\{-\frac{1}{2\sigma^2}[\sum(x_i - 2)^2 - \sum(x_i - 1)^2]\}$$

$$\frac{f_1(\underline{x})}{f_o(\underline{x})} = exp\{-\frac{1}{2\sigma^2}[-2\sum x_i + 3n]\}$$

$$\frac{f_1(\underline{x})}{f_o(\underline{x})} = exp\{\frac{\sum x_i}{\sigma^2} - \frac{3n}{2\sigma^2}\}$$

La regla de decisión que minimiza al error promedio $a\alpha(\delta) + b\beta(\delta)$ consiste en rechazar H_o si

$$\frac{f_1(\underline{x})}{f_o(\underline{x})} > \frac{a}{b}$$

es decir: $\bar{x} > \frac{3}{2} + \frac{\sigma^2}{n} ln(\frac{a}{b})$

Si $\sigma^2 = 2$ y $n = 20$, la región crítica $\mathcal{R}$, que es de la forma $\{\bar{x} > c\}$ depende de a y b:

si $a = b$, $c = 3/2$, pero si $a > b$, $c > 3/2$ y si $a < b$, $c < 3/2$.

En particular, si $a = 2/3$ y $b = 1/3$, $\mathcal{R} = \{\bar{x} > 1{,}57\}$, pero si $a = 1/3$ y $b = 2/3$, $\mathcal{R} = \{\bar{x} > 0{,}143\}$.

El error de tipo I $\alpha(\delta)$ es $\pi(1) = I\!P(\bar{x} > C/\mu = 1)$. Como $\bar{x} \sim \mathcal{N}(1, \sigma^2/n)$ bajo H_o, $\alpha(\delta) = 1 - \Phi(\frac{c-1}{\sigma/\sqrt{n}})$, en que $\Phi(x)$ es la función de distribución de $\mathcal{N}(0, 1)$.

El error $\beta(\delta)$ de tipo II es $1 - \pi(2) = 1 - I\!P(\bar{x} > c/\mu = 1) = I\!P(\bar{x} < c/\mu = 2) = \Phi(\frac{c-2}{\sigma/\sqrt{n}})$

Si $a = b$, como $c = 3/2$, para $n = 20$, se obtiene $\alpha(\delta) = \beta(\delta) = 1 - \Phi(1{,}58) = 0{,}057$, pero con $n = 100$, $\alpha(\delta) = \beta(\delta) = 1 - \Phi(3{,}53) \simeq 0$.

Si se obtuvo una media muestral $\bar{x} = 1{,}30$ para una muestra aleatoria de tamaño 20, no se rechaza $H_o : \mu = 1$ con un error de tipo I de 0.057 cuando se toma a=b; si se toma $a = 0{,}3$ y $b = 0{,}7$, se rechaza H_o a favor de H_1 con un error de tipo I igual a 0.11.

Si tenemos un nivel de significación fijado en $\alpha_o = 0{,}05$, entonces se obtiene una región crítica $\mathcal{R} = \{\bar{x} > c\}$ tal que

$$I\!P(\bar{x} > c|\mu = 1) = 0{,}05$$

Como $\sqrt{n}(\bar{x} - 1) \sim \mathcal{N}(0, 1)$

$$I\!P(\bar{x} > c|\mu = 1) = 1 - \Phi(\sqrt{n}(c - 1)/\sqrt{2}) = 0{,}05$$

Como $\Phi(1{,}65) = 0{,}95$, se obtiene que $\sqrt{n}(c - 1)/\sqrt{2} = 1{,}65$, es decir que $c = 1{,}52$ y $\mathcal{R} = \{\bar{x} > 1{,}52\}$. En este caso no se rechaza H_o.

5.3.3. Tests Uniformemente Más Potentes (U.M.P.)

Vamos extender ahora los resultados del lema de Neyman-Pearson a hipótesis compuestas.

Sean las hipótesis compuestas $H_o : \theta \in \Omega_o$ contra $H_1 : \theta \in \Omega_1$ $(\Omega_o \bigcap \Omega_1 = \emptyset)$.

Si nos fijamos un nivel de significación α_o, buscamos una regla de decisión δ tal que la función de potencia cumpla:

$$\pi_\delta(\theta) \leq \alpha_o \ \forall\theta \in \Omega_o \quad y \quad \pi_\delta(\theta) \text{ sea máxima } \forall\theta \in \Omega_1.$$

Ahora bien no es siempre posible encontrar un test δ que satisfaga esta condición. En efecto si $\Omega_1 = \{\theta_1, \theta_2\}$, un test δ podrá tener una potencia máxima para θ_1 pero no necesariamente para θ_2.

En el ejemplo anterior, si tomamos como hipótesis alternativa dos valores $H_1 = \mu \in \{0, 2\}$, entonces como $H_o : \mu = 1$, la región crítica más potente para $\mu = 0$ sera de la forma $\mathcal{R} = \{\bar{x} < c\}$, que, como lo vimos, no es la región crítica más potente para $\mu = 2$.

Definición 5.3.7 *Si un test δ maximiza la función de potencia para todo valor θ de la hipótesis alternativa $H_1 : \theta \in \Omega_1$, se dice que el test δ es uniformemente más potente (U.M.P.); es decir que δ^* es un test U.M.P. al nivel de significación α_o si $\alpha(\delta) \leq \alpha_o$ y si para todo otro test δ tal que $\alpha(\delta) \leq \alpha_o$, se tiene $\pi_\delta(\theta) \leq \pi_{\delta^*}(\theta)\ \forall\theta \in \Omega_1$.*

Observamos en el ejemplo que la razón de las verosimilitud dado $\mu = \mu_2$ y $\mu = \mu_1$ se escribe:

$$\frac{f_n(\underline{x}|\mu_2)}{f_n(\underline{x}|\mu_1)} = exp\{\frac{n(\mu_2 - \mu_1)}{\sigma^2}(\bar{x} - \frac{1}{2}(\mu_2 - \mu_1))\}$$

Se observa que $\dfrac{f_n(\underline{x}|\mu_2)}{f_n(\underline{x}|\mu_1)}$ depende de $\underline{x}$ sólo a través de la media muestral $\bar{x}$; además crece en función de $\bar{x}$ si $\mu_1 < \mu_2$. Este cociente es monótono con respecto a $\bar{x}$.

Definición 5.3.8 *Se dice que $f_n(\underline{x}/\theta)$ tiene una razón de verosimilitud monótona para un estadístico g($\underline{x}$) si y sólo si $\forall\theta_1, \theta_2$ tal que $\theta_1 < \theta_2$, el cociente $\dfrac{f_n(\underline{x}|\theta_2)}{f_n(\underline{x}|\theta_1)}$ depende del vector $\underline{x}$ a través de la función $g(\underline{x})$ y el cociente es una función creciente de $g(\underline{x})$ $\forall\underline{x}$.*

En el ejemplo anterior $f_n(\underline{x}|\mu)$ tiene una razón de verosimilitud monótona en $\underline{x}$.

Veamos el ejemplo de una Bernoulli de parámetro p, tomando $y = \sum x_i$, $f_n(\underline{x}|p) = p^y(1 - p)^{n-y}$. Si $0 < p_1 < p_2 < 1$

$$\frac{f_n(\underline{x}|p_2)}{f_n(\underline{x}|p_1)} = \frac{(p_2(1 - p_1))^y}{(p_1(1 - p_2))^y}\frac{(1 - p_2)^n}{(1 - p_1)^n}$$

cociente que depende de $\underline{x}$ a través de y. Además es una función creciente de y y por lo tanto tiene una razón de verosimilitud monótona en $\sum x_i$.

Definición 5.3.9 *Un test sobre las hipótesis $H_o : \theta \leq \theta_o$ contra $H_1 : \theta > \theta_o$, se dice unilateral y un test sobre las hipótesis $H_o : \theta = \theta_o$ contra $H_1 : \theta \neq \theta_o$, se dice bilateral.*

Vamos a mostrar que si $f_n(\underline{x}|\theta)$ tiene una razón de verosimilitud monótona en algún estadístico T, entonces existe un test U.M.P. para las hipótesis $H_o : \theta \leq \theta_o$ contra $H_1 : \theta > \theta_o$

Teorema 5.3.10 *Si $f_n(\underline{x}/\theta)$ tiene una razón de verosimilitud monótona en el estadístico T y si c es la constante tal que $I\!P(T \geq c|\theta = \theta_o) = \alpha_o$, entonces la regla de decisión que permite rechazar la hipótesis nula si $T \geq c$ es un test U.M.P. para $H_o : \theta \leq \theta_o$ contra $H_1 : \theta > \theta_o$ al nivel de significación α_o.*

Demostración Sea θ_1 tal que $\theta_1 > \theta_o$, $\alpha(\delta) = I\!P($ rechazar $H_o|\theta = \theta_o) = \pi_\delta(\theta_o)$ y $\beta(\delta) = I\!P($ aceptar $H_o|\theta = \theta_1) = 1 - \pi_\delta(\theta_1)$.

Del lema de Neyman-Pearson, se deduce que entre todos los procedimientos δ de error de tipo I $\alpha(\delta) < \alpha_o$, el valor de $\beta(\delta)$ será mínimo para el procedimiento δ^* que consiste en rechazar H_o cuando $\dfrac{f_n(\underline{x}|\theta_1)}{f_n(\underline{x}|\theta_o)} \geq k$, k siendo elegido de tal forma que $I\!P($rechaza $H_o|\theta = \theta_o) \leq \alpha_o$.

Como $\frac{f_n(\underline{x}|\theta_1)}{f_n(\underline{x}|\theta_o)}$ es una función creciente de T, un procedimiento, que rechaza H_o cuando el cociente es al menos igual a k, es equivalente al procedimiento que rechaza H_o cuando T es al menos igual a una constante c.

La constante c se elige de tal forma que $I\!P($rechazar $H_o/\theta = \theta_o) \leq \alpha_o$.

Ahora bien esto es cierto para todo $\theta_1 > \theta_o$ y por lo tanto este procedimiento es U. M. P. para $H_o : \theta = \theta_o$ contra $H_1 : \theta > \theta_o$.

Por otro lado, la función de potencia es no decreciente en θ y por lo tanto si $\pi(\theta_o|\delta) \leq \alpha_o$, entonces $\pi(\theta|\delta) \leq \alpha_o \ \forall\theta \leq \theta_o$. ■

Cuando $f_n(\underline{x}|\theta)$ no tiene una razón de verosimilitud monótona, el test de razón de verosimilitud permite resolver una gran cantidad de problemas:

Si $H_o : \theta \in \Omega_o$ contra $H_1 : \theta \in \Omega_1$, se define

$$\lambda(\underline{x}) = \frac{Sup f_n(\underline{x}|\theta \in \Omega_1)}{Sup f_n(\underline{x}|\theta \in \Omega_o)}$$

El test de razón de verosimilitud consiste en rechazar H_o si $\lambda(\underline{x}) > k$ y en no rechazar H_o si $\lambda(\underline{x}) < k$.

El problema es encontrar entonces la distribución de $\lambda(\underline{x})$. El siguiente teorema da una solución aproximada a este problema.

Teorema 5.3.11 *Si θ es un vector de parámetros de dimensión r y si la hipótesis nula es función lineal de θ, $H_o : H\theta = 0$, en que $H \in \mathcal{M}_{sxr}$, entonces $-2ln\lambda(\underline{x})$ tiene una distribución asintótica χ^2_s.*

5.4. TESTS PARAMÉTRICOS USUALES

Veamos algunos tests usuales que se basan en los resultados anteriores.

5.4.1. Test sobre una media con la varianza conocida

Sea una v.a. $X \sim \mathcal{N}(\mu, \sigma^2)$ en que la varianza σ^2 es conocida e igual a 36^2 y una muestra aleatoria de tamaño $n = 9$.

Sean las hipótesis $H_o : \mu = 180$ contra $H_1 : \mu > 180$ y un nivel de significación de $\alpha = 0{,}05$.

De lo anterior, se deduce que la región crítica más potente es de la forma $\mathcal{R} = \{\bar{x} > c\}$ con c determinado por:

$$I\!P(\bar{x} \geq c | \mu = 180) = 0{,}05$$

Como $\bar{x} \sim \mathcal{N}(\mu, 144)$, $(\bar{x} - \mu)/12 \sim \mathcal{N}(0, 1)$.

Si la hipótesis nula $H_o : \mu = 180$ es cierta $(\bar{x} - 180)/12 \sim \mathcal{N}(0, 1)$. Luego la constante c se determina de:

$$I\!P(\bar{x} - 180)/12 \geq (c - 180)/12) = 0{,}05$$

Utilizando las tablas estadísticas se obtiene $(c - 180)/12 = 1{,}65$. Finalmente $c = 200$.

La región crítica $\{\bar{x} \geq 200\}$ es U. M. P. para todo $\mu > 180$ de la hipótesis alternativa.

El error de tipo II depende de μ. Como lo muestra la tabla 5.1 y el gráfico 5.2 , el error de tipo II aumenta cuando el valor de μ es muy cercano al valor 180 de H_o.

μ	180	185	190	200	210	220	230
$\pi(\mu)$	0.05	0.11	0.20	0.50	0.80	0.95	0.994
$1 - \pi(\mu)$	0.95	0.89	0.80	0.50	0.20	0.05	0.006

Cuadro 5.1: Potencia y error de tipo II para $H_1 : \mu > 180$

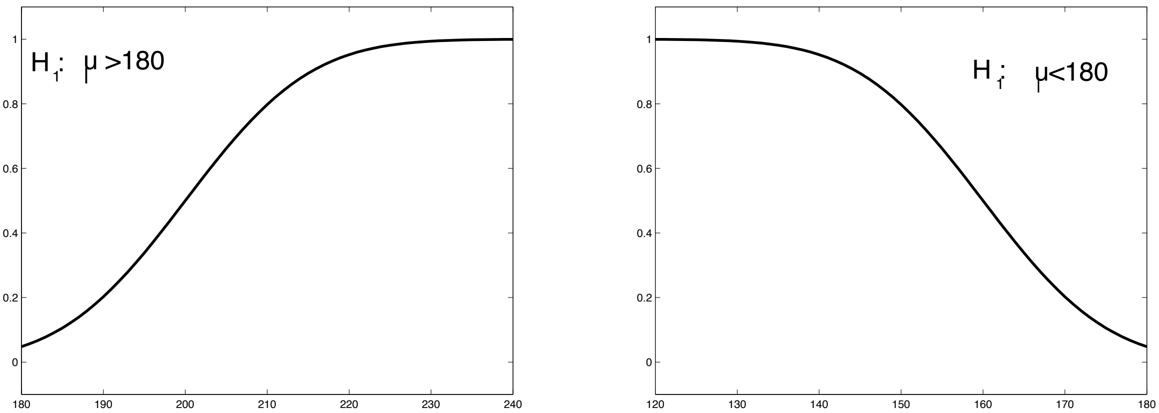

Figura 5.2: Función de Potencia para $H_1 : \mu > 180$ y para $H_1 : \mu < 180$

Sea ahora $H_o : \mu = 180$ contra $H_1 : \mu < 180$ con un nivel de significación de 0.05.

La región crítica más potente es de la forma $\mathcal{R} = \{\bar{x} \leq c\}$ con c determinado por:

$$I\!P(\bar{x} \leq c | \mu = 180) = 0{,}05$$

La región crítica $\{\bar{x} \leq 160\}$ es U. M. P. para todo $\mu < 180$ de la hipótesis alternativa. La función de potencia esta dada en la tabla 5.2 y el gráfico 5.2.

μ	180	175	170	160	150	140	130
$\pi(\mu)$	0.05	0.11	0.20	0.50	0.80	0.95	0.99
$1 - \pi(\mu)$	0.95	0.89	0.80	0.50	0.20	0.05	0.006

Cuadro 5.2: Potencia y error de tipo II para $H_1 : \mu < 180$

Sea finalmente $H_o : \mu = 180$ contra $H_1 : \mu \neq 180$ con un nivel de significación de 0.05.

No existe un test U. M. P. para est hipótesis alternativa; se propone como región crítica

$$\mathcal{R} = \{\bar{x} \leq a\} \cup \{\bar{x} \geq b\}$$

de tal forma que $I\!P(\bar{x} \leq a) = 0{,}025$ y $I\!P(\bar{x} \geq b) = 0{,}025$. Obtenemos $a = 156{,}5$ y $b = 203{,}5$, que produce una función de potencia presentada en la tabla 5.3 y el gráfico 5.3. Se nota que la potencia es siempre inferior o igual a la potencia de la tabla 5.1 ó 5.2 para todo μ.

μ	140	150	160	170	175	180	185	190	200	210	220
$\pi(\mu)$	0.91	0.69	0.37	0.12	0.07	0.05	0.07	0.12	0.37	0.69	0.91
$1 - \pi(\mu)$	0.09	0.31	0.43	0.88	0.93	0.95	0.93	0.88	0.43	0.31	0.09

Cuadro 5.3: Función de Potencia para $H_1 : \mu \neq 180$

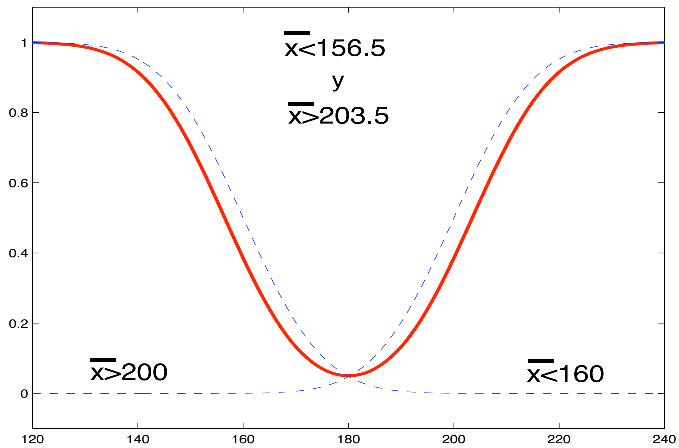

Figura 5.3: Función de Potencia para $H_1 : \mu \neq 180$

Se observara que este test se basa en el supuesto de distribución normal de los valores muestrales. Cuando el tamaño de la muestra es grande, este supuesto es aceptable, pero para muestras pequeñas, es importante comprobar si lo es.

5.4.2. Test sobre una media con la varianza desconocida

Si en el problema anterior suponemos que la varianza es desconocida, se procede de manera parecida al caso de la varianza conocida utilizando la distribución de t Student de la variable $\frac{(\bar{x}-\mu)}{s_n/\sqrt{n-1}}$ que es una Student a $n-1$ grados de libertad.

5.4.3. Test sobre una varianza

Si ahora planteamos las hipótesis:

$$H_o : \sigma^2 \geq \sigma_o^2 \qquad contra \qquad H_1 : \sigma^2 < \sigma_o^2,$$

en donde σ_o^2 es un escalar positivo dado.

A partir del estadístico $\frac{ns_n^2}{\sigma_o^2}$, que sigue una distribución de χ^2 a $n-1$ grados de libertad bajo H_o, se construye la región critica de nivel de significación α:

$$\mathbb{P}(\frac{ns_n^2}{\sigma_o^2} \leq c) = \alpha$$

5.4.4. Test de comparación de dos medias

Frecuentemente se está interesado en hacer inferencia sobre la diferencia de dos medias. Por ejemplo, la diferencia de sueldos promedios μ_1 y μ_2 entre dos poblaciones Ω_1 y Ω_2 o la eficiencia de un tratamiento comparando los resultados de la población que utilizó el tratamiento con la que no lo utilizó. Las hipótesis se escriben entonces:

$$H_o : \mu_1 - \mu_2 = d_o$$

$$H_1 : \mu_1 - \mu_2 \neq d_o$$

Es usual tomar $d_o = 0$ y la hipótesis alternativa H_1 puede ser

$$H_1 : \mu_1 - \mu_2 \neq 0 \qquad o \qquad H_1 : \mu_1 - \mu_2 > 0$$

Definamos la v.a. $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ y X $\sim \mathcal{N}(\mu_2, \sigma_2^2)$ como los sueldos de los individuos en las poblaciones Ω_1 y Ω_2 respectivamente. Si la media muestral de X obtenida sobre una muestra de tamaño n_1 en Ω_1 es $\bar{x}_1$ y la media muestral obtenida sobre una muestra de tamaño n_2 en Ω_2 es $\bar{x}_2$, entonces

$$\bar{x}_1 - \bar{x}_2 \sim \mathcal{N}(\mu_1 - \mu_2, \sigma_1^2/n_1 + \sigma_2^2/n_2)$$

Si las varianzas σ_1^2 y σ_2^2 son conocidas, entonces se obtiene una región crítica de nivel de significación $\alpha = 0{,}05$ para $H_o : \mu_1 = \mu_2$ contra $H_1 : \mu_1 > \mu_2$:

$$\mathbb{P}(\bar{x}_1 - \bar{x}_2 \geq 1{,}96\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2})$$

Si las varianzas son desconocidas, pero se suponen iguales ($\sigma_1^2 = \sigma_2^2 = \sigma^2$), entonces se estima esta varianza y se usa un estadístico que sigue una distribución t de Student. Si s_1^2 y s_2^2 son las varianzas empíricas sesgadas para σ_1^2 y σ_2^2 respectivamente, el estimador de σ^2 es:

$$s^2 = (n_1 s_1^2 + n_2 s_2^2)/(n_1 + n_2 - 2)$$

es insesgado para σ^2. Entonces

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2})(\frac{n_1 + n_2}{n_1 n_2})}}$$

es una Student a $n_1 + n_2 - 2$ grados de libertad.

La región crítica se define entonces como:

$$I\!\!P(\bar{x}_1 - \bar{x}_2 \geq t_\alpha \sqrt{(\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2})(\frac{n_1 + n_2}{n_1 n_2})})$$

en donde t_α es tal que $I\!\!P(t_{n_1+n_2-2} \geq t_\alpha) = \alpha$.

Aquí se hizo el supuesto de la igualdad de las varianzas y de la independencia de las dos muestras.

5.4.5. Test para pares de observaciones

Hay situaciones en donde las muestras no son independientes. Es el caso cuando se toman muestras formadas de pares, es decir cuando cada observación de una muestra está relacionada a una observación de la otra muestra. Por ejemplo, se mide la eficacia de un tratamiento comparando un exámen médico sobre la misma población pero antes y después del tratamiento, o bien se considera la diferencia de edad en un grupo de parejas donde una muestra esta formada por las esposas y la otra muestra por sus correspondientes maridos. La dos muestras no son independientes y son del mismo tamaño. Sean (X, Y) las v.a. edades de la mujer y su marido y una muestra de n matrimonios $\{(x_i, y_i), i = 1, 2, ..., n\}$. La diferencia entre las medias empíricas $\bar{x}_n$ e $\bar{y}_n$ es un estimador insesgado de la diferencia $\mu_1 - \mu_2$ entre las dos poblaciones apareadas:

$$E(\bar{x}_n - \bar{y}_n) = E(X - Y) = E(X) - E(Y) = \mu_1 - \mu_2$$

Pero debido a la dependencia entre X e Y la varianza de la diferencia X-Y cambia.

$$\sigma_{X-Y}^2 = E(X - Y - (\mu_1 - \mu_2))^2) = E(X - \mu_1)^2 + E(Y - \mu_2)^2 - 2E(X - \mu_1)(Y - \mu_2)$$

$$\sigma_{X-Y}^2 = \sigma_1^2 + \sigma_2^2 - 2Cov(X, Y)$$

Como no se conoce en general las varianzas σ_1^2, σ_2^2 y la covarianza $Cov(X, Y)$, se estima la varianza de la diferencia σ_{X-Y}^2 considerando que los valores muestrales son las diferencias $d_i = x_i - y_i$ que provienen de una sola muestra:

$$\hat{\sigma}_{X-Y}^2 = \frac{\sum (d_i - \bar{d}_n)^2}{n}$$

en donde $\bar{d}_n = \frac{\sum d_i}{n} = \frac{\sum x_i - y_i}{n}$. Se puede escribir:

$$\hat{\sigma}^2_{X-Y} = \frac{\sum (x_i - y_i)^2}{n} - \bar{d}^2_n$$

El estimador de la varianza de $\bar{d}_n$ es igual entonces a $\frac{\hat{\sigma}^2_{X-Y}}{n}$ y $\frac{\bar{x}_n - \bar{y}_n - (\mu_1 - \mu_2)}{\hat{\sigma}_{X-Y}/\sqrt{n-1}}$ sigue una t de Student a $n-1$ grados de libertad.

5.4.6. Test de comparación de dos varianzas: la distribución F

Se quiere comparar las varianzas σ_1^2 y σ_2^2 de dos poblaciones normales a partir de muestras aleatorias independientes de cada población. Si $x_1, x_2, ..., x_n$ e $y_1, y_2, ..., y_m$ son las muestras aleatorias respectivas, $s_1^2 = \frac{1}{n}\sum (x_i - \bar{x})^2$ y $s_2^2 = \frac{1}{n}\sum (y_i - \bar{y})^2$ son las varianzas muestrales, $U = ns_1^2/\sigma_1^2 \sim \chi^2_{n-1}$ y $V = ms_2^2/\sigma_2^2 \sim \chi^2_{m-1}$; además U y V son independientes.

Vimos en el capítulo anterior que $\frac{U/(n-1)}{V/(m-1)}$ sigue una distribución F de Fisher a $n-1$ y $m-1$ grados de libertad.

Consideremos entonces el estadístico

$$\frac{ns_1^2/(n-1)}{ms_2^2/(m-1)}$$

que sigue una distribución $F_{n-1,m-1}$ bajo la hipótesis nula $H_o : \sigma_1 = \sigma_2$.

Se define la región crítica de nivel de significación α para $H_o : \sigma_1^2 = \sigma_2^2$ contra $H_1 : \sigma_1^2 > \sigma_2^2$ como:

$$I\!P\left(\frac{nsS_1^2/(n-1)}{ms_2^2/(m-1)} > F_\alpha\right) = \alpha$$

en donde F_α se calcula a partir de la F de Fisher a $n-1$ y $m-1$ g.l.

5.5. TESTS χ^2

Se habla de test χ^2 a todo test que use un estadístico que sigue una distribución χ^2. Aquí trataremos los tests χ^2 obtenidos a partir de una distribución multinomial. Veremos previamente dos distribuciones de vectores aleatorios, la distribución normal multivariada y la distribución multinomial que tiene un comportamiento asintótico de vector normal multivariado. Después de presentar un test para un modelo multinomial, veremos aplicaciones del test para hipótesis no paramétricas.

5.5.1. La distribución normal multivariada

Se puede definir de dos maneras la distribución normal multivariada.

Sea $X = \begin{pmatrix} X_1 \\ X_2 \\ . \\ . \\ X_p \end{pmatrix}$ un vector aleatorio de $I\!R^p$.

Definición 5.5.1 *Sea u: $I\!R^p \longrightarrow I\!R$*
Se dice que X es un vector normal multivariado de orden p de vector de media μ y de matriz de varianza-covarianza Γ (se denota $X \sim \mathcal{N}_p(\mu,\Gamma)$) si y sólo sí $u(X) \sim \mathcal{N}(u(\mu),\Gamma(u,u))$.

Es decir que si X es un vector normal, toda combinación lineal de X es una v.a. normal.

Definición 5.5.2 *Se dice que $X \sim \mathcal{N}_p(\mu,\Gamma)$ si su función característica es*

$$\Psi_X(u) = exp(iu^t\mu - \frac{1}{2}u^t\Gamma u) \qquad \forall u \in I\!R^p$$

Propiedades:

- Tomando como vector u los vectores canónicos, se obtienen las leyes marginales de X, que son normales; pero la recíproca es falsa: un vector formado de variables normales no es necesariamente un vector normal: En efecto Γ tiene que ser semidefinida positiva ya que $\Gamma(u) = u^t\Gamma u$ es la varianza de u^tX.
- Sea Y una matriz $(p \times q)$.
$$X \sim \mathcal{N}_p(\mu,\Gamma) \Longrightarrow Y = AX \sim \mathcal{N}_q(A\mu, A\Gamma A^t).$$
- Las v.a. X_i son independientes $\Longleftrightarrow$ Γ es diagonal
- Γ es semidefinida positiva
En efecto $\Gamma(u,u) = u^t\Gamma u$ es la varianza de la v.a. $u(X) = u^tX$.
- Si Γ es de rango r, existe Λ una matriz $(r \times p)$ tal que $\Gamma = \Lambda\Lambda^t$. Entonces:
$$X \sim \mathcal{N}_p(\mu,\Gamma) \Longleftrightarrow X = \mu + \Lambda Y \quad Y \sim \mathcal{N}_r(0, I_r)$$
es decir que las componentes del vector Y son centradas, normalizadas y independientes entre si.
- Si Γ es invertible, Λ es invertible también e $Y = \Lambda^{-1}(X-\mu)$.

Este último resultado permite calcular la densidad del vector X. En efecto se puede calcular la densidad del vector $Y \sim \mathcal{N}_p(0, I_p)$:

$$f(Y) = \prod f(Y_i) = (\frac{1}{2\pi})^{p/2} exp(-\frac{1}{2}\sum Y_i^2)) = (1/2\pi)^{p/2} exp(-\frac{1}{2}YY^t)$$

Como $YY^t = (\Lambda^{-1}(X-\mu)^t\Lambda^{-1}(X-\mu) = (X-\mu)^t\Gamma^{-1}(X-\mu)$, el Jacobiano de la transformación es $|\Gamma|^{-1/2}$, luego la densidad de X es:

$$h(X) = (\frac{|\Gamma|^{-1/2}}{2\pi})^{p/2} exp(-\frac{1}{2}(X-\mu)^t\Gamma^{-1}(X-\mu))$$

Proposición 5.5.3 *Si el vector aleatorio* $X \sim \mathcal{N}(\mu, \Gamma)$ *con* Γ *de rango* r*, entonces* $\| X - \mu \|^2_{\Gamma^{-1}} \sim \chi^2_r$.

Demostración Acordamos que si $Y \sim \mathcal{N}(0, I_r)$, $\| Y \|^2 = \sum Y_i^2 \sim \chi^2_r$. Como $\Gamma = \Lambda\Lambda^t$, existe Y tal que $X = \mu + \Lambda Y$, con $Y \sim \mathcal{N}(0, I_r)$. Pero se puede escribir $Y = (\Lambda^t\Lambda)^{-1}\Lambda^t(X - \mu)$, luego:

$$\| Y \|^2_{I_p} = YY^t = \| X - \mu \|^2_{\Gamma^{-1}} \sim \chi^2_r$$

■

5.5.2. La distribución multinomial

La distribución multinomial es una generalización de la distribución binomial. En vez de tener dos alternativas en cada experimento, se tienen k alternativas ($k \geq 2$). Por ejemplo, hay seis resultados posibles cuando se tira un dado. Si el "1" tiene probabilidad p_1, el "2" tiene probabilidad p_2,...,el "6" tiene probabilidad p_6, y si hacemos n lanzamientos independientes, los números M_1 de "1", M_2 de "2",..., M_6 de "6" constituyen un vector aleatorio M con una distribución multinonimal de parámetros $n, p_1, p_2, ..., p_6$. Se observa que $\sum M_i = n$ y

$$I\!P(M = m) = I\!P(M_1 = m_1, ..., M_6 = m_6) = \frac{n! p_1^{m_1} p_2^{m_2} ... p_6^{m_6}}{m_1! m_2! ... m_6!}$$

Calculemos la esperanza y la varianza de M.

Si $p = \begin{pmatrix} p_1 \\ p_2 \\ . \\ p_6 \end{pmatrix}$, entonces $E(M) = np$.

Sea el resultado J_i del lanzamiento i: $J_i = e_h$, el h-ésimo vector canónico si el resultado es h. Entonces $M = \sum J_i$.

$E(J_i) = p$ y $E(J_i J_i^t) = \sum_h e_h e_h^t I\!P(J_i = e_h) = \begin{pmatrix} p_1 & 0 & . & 0 \\ 0 & p_2 & . & 0 \\ . & . & . & . \\ 0 & 0 & . & p_6 \end{pmatrix} = Diag(p)$

$Var(J_i) = E(J_i J_i^t) - E(J_i)[E(J_i)]^t = Diag(p) - pp^t = \Sigma(p)$

Luego $Var(M) = n\Sigma(p)$.

Por el Teorema del Límite Central, se tiene:

$$\lim_{n \longrightarrow +\infty} I\!P(\frac{M - np}{\sqrt{n}} \leq x) = \Phi(x),$$

en donde Φ es la función de distribución normal multivariada centrada de matriz de covarianza $\Sigma(p)$.

Ejercicio: Muestre que si el vector multinomial es de dimensión k, entonces el rango de la matriz $\Sigma(p)$ es igual a $k - 1$ (Se puede mostrar que el núcleo de $\Sigma(p)$ es de dimensión 1).

Proposición 5.5.4 *Si M es un vector de distribución multinomial $\mathcal{M}(n, p_1, ..., p_k)$, entonces*

$$Q = \sum \frac{(M_i - np_i)^2}{np_i}$$

tiene una distribución asintótica de χ^2_{k-1}.

La demostración se basa en el resultado del ejercicio anterior.

5.5.3. Test de ajuste para un modelo multinomial

Sea un dado que se tira $n = 102$ veces. Se obtiene entonces la distribución empírica (tabla 5.4).

¿Podemos concluir que el dado esta cargado?

Sea la hipótesis nula $H_o : p_i = \frac{1}{6} \quad \forall i$ que el dado es equilibrado.

Entonces, si la hipótesis nula es cierta, $M \sim \mathcal{M}(102, 1/6, ..., 1/6)$.

Calculamos entonces el estadístico Q para construir la región crítica (tabla 5.5).

i	1	2	3	4	5	6	Total
M_i	12	11	22	20	16	21	102

Cuadro 5.4: Resultados del experimento

i	M_i	np_i	$M_i - np_i$	$(M_i - np_i)^2/np_i$
1	12	17	-5	1.471
2	11	17	-6	2.118
3	22	17	5	1.471
4	20	17	3	0.529
5	16	17	-1	0.059
6	21	17	4	0.941
Total	102	102	0	6.589

Cuadro 5.5: Calculo de Q

Se obtiene $Q = 6{,}589$, y $I\!P(\chi^2_5 > 6{,}589) > 5\,\%$, por lo cual no se rechaza H_o. Lo que quiere decir que entre los valores observados y los valores esperados las diferencias no son suficientemente significativas como para decidir que el dado esta cargado.

5.5.4. Test de ajuste para una distribución discreta

Se considera el número de accidentes X observados cada fin de semana en una carretera peligrosa (tabla 5.6). Se quiere probar la hipótesis que X sigue una distribución de Poisson de parámetro λ a partir de datos obtenidos sobre un año. En primera instancia supondremos λ conocido e igual a 1,5. Se tiene entonces $H_o : X \sim \mathcal{P}(1{,}5)$.

N^o accidentes	0	1	2	3	4	5	6 y más	Total
N^o semanas	17	16	10	5	2	1	1	52

Cuadro 5.6: Resultados del experimento

Bajo H_o, los números de semanas M_o con "0" accidente, M_1 con "1" accidente, ..., M_6 con "6" o más accidentes sigue una distribución multinomial de parámetros $n = 52$, y $p_o = I\!\!P(X = 0)$, $p_1 = I\!\!P(X = 1)$, ..., $p_6 = I\!\!P(X \geq 6)$,

Calculemos los $p_i = I\!\!P(X = i)$ bajo el supuesto que $X \sim \mathcal{P}(1{,}5)$ y comparemos los valores np_i con los valores observados M_i (tabla 5.7).

i	M_i	np_i	$M_i - np_i$	$(M_i - np_i)^2/np_i$
0	17	11.60	5.400	2.5124
1	16	17.40	0.596	0.0204
2	10	13.05	-3.052	0.7137
3	5	6.53	-1.526	0.3568
4	2	2.45	-0.449	0.0824
5	1	0.73	0.267	0.0971
6	1	0.24	0.766	3.2735
Total	52	52	0	7.0563

Cuadro 5.7: Calculo de Q

Obtienemos $Q = 7{,}0563$, y $I\!\!P(\chi^2_6 > 7{,}0563) > 5\,\%$, por lo cual no se rechaza H_o.

Ahora si se supone que no se conoce el parámetro λ, se puede estimar por $\hat{\lambda} = \bar{X}_n = \sum iM_i/52 = 72/52 = 1{,}385$ y proceder como antes. Pero ahora el estadístico Q pierde un grado de libertad debido a la estimación.

Con el parámetro $\hat{\lambda}$, Q=5.62 y $I\!\!P(\chi^2_5 > 5{,}62) > 5\,\%$.

5.5.5. Test de ajuste para una distribución continua

Para construir un test χ^2 para una hipótesis relativa a una distribución continua, por ejemplo $H_o : X \sim \mathcal{N}(1, 0{,}25)$, basta transformar la variable en una variable discreta. Se divide el rango de X en k intervalos disjuntos I_1, I_2, ...,I_k y se cuentan los números de observaciones de la muestra M_i que caen en el intervalo I_i. El vector M formado de los M_i sigue una distribución multinomial de parámetros de probabilidad determinados por la hipótesis nula.

Sea por ejemplo, las temperaturas medias X del mes de septiembre en la Urbe durante 60 años (tabla 5.8). Se quiere probar la hipótesis nula $H_o : X \sim normal$.

Hay diferentes maneras de definir la partición de intervalos de $I\!\!R$. Una vez fijado el número de intervalos, se pueden elegir del mismo largo o de la misma probabilidad. Tomaremos aquí 10 intervalos equiprobables.

Para calcular las probabilidades, hay que estimar previamente los parámetros μ y σ^2 de la normal:

$$\hat{\mu} = \bar{x}_n = 15{,}76 \quad \hat{\sigma}^2 = s_n^2 = 13{,}82$$

Luego los intervalos I_j se obtienen (tabla 5.9) de tal forma que:

$$\mathbb{P}(X \in I_j) = 0{,}10 \quad \forall j$$

en donde $X \sim \mathcal{N}(15{,}76, 13{,}82)$.

Evaluando se obtiene $Q = 9{,}35$. El estadístico χ^2 tiene aquí 7 grados de libertad. (Se estimaron dos parámetros). Como $\mathbb{P}(\chi^2_7 > 9{,}35) > 5\,\%$, no se rechaza la hipótesis de normalidad.

5.2	6.5	7.5	8.2	10.1	10.5	11.6	12.0	12.0	12.8	13.5	13.8
13.9	14.0	14.0	14.2	14.3	14.5	14.7	14.8	15.0	15.0	15.2	15.2
15.3	15.4	15.6	15.8	15.8	15.9	16.0	16.1	16.2	16.4	16.4	16.5
16.5	16.8	16.9	17.0	17.0	17.1	17.1	17.1	17.4	17.6	17.9	18.2
18.5	18.8	18.9	19.4	19.8	20.3	20.9	21.4	21.9	22.5	22.8	23.9

Cuadro 5.8: Temperaturas medias

I_i	M_i	np_i	$M_i - np_i$	$(M_i - np_i)^2/np_i$
$]-\infty$,10.96]	6	6	0	0.00
]10.96,12.64]	3	6	-3	1.50
]12.64,13.83]	3	6	-3	1.50
]13.83,14.83]	8	6	2	0.67
]14.83,15.76]	7	6	1	0.17
]15.76,16.69]	10	6	4	2.67
]16.69,17.69]	9	6	3	1.50
]17.69,18.88]	4	6	-2	0.67
]18.88,20.56]	4	6	-2	0.67
]20.56,$+\infty$]	6	6	0	0.00
Total	60	60	0	9.35

Cuadro 5.9: Calculo de Q

5.5.6. Test de independencia para 2 variables nominales

Cuando dos v.a. discretas con valores en conjuntos A y B respectivamente son independientes:

$$\mathbb{P}(X = i \;\; e \;\; Y = j) = \mathbb{P}(X = i)\mathbb{P}(Y = j) \quad \forall (i,j) \in A \times B$$

Si A y B son conjuntos finitos ($card(A) = p$, $card(B) = q$), las frecuencias M_{ij} de observaciones obtenidas en una muestra bivariada de tamaño n siguen una distribución multinomial de parámetros n, p en donde p es el vector de las probabilidades $p_{ij} = \mathbb{P}(X = i \;\; e \;\; Y = j)$.

Bajo la hipótesis de independencia entre X e Y, se puede estimar estos parámetros p_{ij} a partir de las frecuencias marginales de X e Y:

$$\hat{p}_{ij} = \hat{p}_{i\bullet}\hat{p}_{\bullet j}$$

con $\hat{p}_{i\bullet} = \sum_j M_{ij}/n$ y $\hat{p}_{\bullet j} = \sum_i M_{ij}/n$.

Lo que permite usar el estadístico

$$Q = \sum_{ij} \frac{(M_{ij} - n\hat{p}_{i\bullet}\hat{p}_{\bullet j})^2}{n\hat{p}_{i\bullet}\hat{p}_{\bullet j}}$$

que sigue una distribución asintótica χ^2 a $(p-1)(q-1)$ grados de libertad.

Sea un conjunto de consumidores que dan su apreciación sobre un chocolate. Se quiere estudiar si existe una relación entre la opinión de los consumidores y su nivel socio-económico (NSE).

Se considera la tabla de contingencia obtenida a partir de une encuesta de estudio de mercado sobre 1600 consumidores (tabla 5.10), que presenta las frecuencias M_{ij} para cada NSE i y apreciación j.

NSE	APRECIACION			TOTAL
	MALA	REGULAR	BUENA	
A	140	100	45	285
B	50	225	350	625
C	15	175	500	690
TOTAL	205	500	895	1600

Cuadro 5.10: Tabla de contingencia

Las probabilidades p_{ij} se estiman usando las frecuencias marginales de la tabla; por ejemplo, para el NSE A con la apreciación MALA se obtiene $\hat{p}_{11} = 285 \times 205/1600 = 0{,}0228$ y $n\hat{p}_{11} = 36{,}51$.

Se obtiene el valor $Q = 521{,}46$. Como $I\!\!P(\chi_4 > 521{,}46) < 5\,\%$, por tanto, se rechaza la hipótesis de independencia entre el NSE y la apreciación.

Nota: Se puede usar el mismo test para probar la independencia de dos variables continuas transformándolas en variables discretas.

5.6. EJERCICOS

1. El cocinero del casino preparó la masa para hacer 500 empanadas. Ese mismo día, en un grupo de 20 alumnos que almorzaron juntos, alguien propuso contar la cantidad de pasas que cada uno encontrase en su empanada, encontrándose la siguiente distribución:

N^o de pasas	0	1	2	3	4	5	8
N^o de empanadas	1	3	4	5	4	2	1

a) Suponiendo que la distribución de la cantidad de pasas X en una empanada sigue una ley de Poisson, estime el parámetro λ de esta ley.
b) Justifique la hipótesis: H_0: La distribución de la cantidad de pasas en una empanada sigue una ley de Poisson de las dos formas siguientes:
(i) A priori: Buscando la probabilidad de que una empanada tenga exactamente x pasas.
(ii) A posteriori: comparando los resultados esperados bajo la hipótesis con aquellos observados en la muestra.
c) Se decide que las empanadas son *aceptables* si en promedio cada empanada tiene 3,5 pasas; el cocinero afirma que está es la cantidad de pasas por empanadas. Los alumnos, en cambio, objetan que las empanadas tienen en promedio sólo 2,5 pasas.
¿Qué significa la elección de los test de hipótesis siguientes:

$$H_0\text{: } \lambda = 3.5 \text{ vs. } H_1\text{: } \lambda = 2.5 \qquad H_0'\text{: } \lambda = 2.5 \text{ vs } H_1'\text{: } \lambda = 3.5 \text{ ?}$$

d) Dé la región crítica del test H_0 vs. H_1 al nivel de significación $\alpha = 0{,}05$. Dé la potencia de este test y concluir si las empanadas son *aceptables*.
e) Misma pregunta tomando H_0' vs H_1'.
f) Comparar las dos decisiones anteriores.

2. Se tienen los pesos de diez parejas antes y después de 6 meses de matrimonio:

	antes	72	69	81	71	88	78	68	76	86	95
Hombres	después	77	68.5	85	74.5	90.5	76	71	75	87.5	101
	antes	52	56	61	49	57	63	66	59	67	51
Mujeres	después	54	55	58	50	55	61	64	56	70	50

¿Cuál es la influencia del matrimonio sobre el peso de los hombres y de las mujeres?

3. Se quiere probar si hay una diferencia de ingreso entre hombres y mujeres médicos. Se hizo una encuesta a $n = 200$ médicos seleccionados al azar e independientemente. Se obtuvo la siguiente información:

	Ingresos bajos	Ingresos altos	Total
Hombres	20	100	120
Mujeres	70	10	80
Total	90	110	200

a) Sean p_1 y p_2 las proporciones poblacionales de médicos hombres y mujeres; y sean p_1' y p_2' las proporciones poblacionales de médicos con ingresos bajos y altos. Realice los tests

$$H_0 : p_1' = p_2 \quad vs. \quad H_1 : p_1' \neq p_2 \qquad H_0' : p_1 = p_2' \quad vs. \quad {}_1' : p_1 \neq p_2'$$

b) Estudie la independencia entre sexo e ingreso.

4. Supóngase que X_1, ..., X_n constituyen una m.a.s. de X con distribución uniforme sobre $[0, \theta]$ y las siguientes hipótesis:

$$H_0 : \theta \geq 2 \quad vs. \quad H_1 : \theta < 2$$

Sea $Y_n = max\{X_1, ..., X_n\}$ y considérese la región critíca $\{Y_n \leq 1{,}5\}$.
a) Determínese la función de potencia.
b) Determínese el tamaño del test.

5. Suponga que se desconoce la proporción p de artículos defectuosos en una población de artículos y se desea probar las hipótesis

$$H_0 : p = 0{,}2 \quad vs \quad H_1 : p \neq 0{,}2$$

Suponga además que se selecciona una m.a.s. de tamaño 20. Sea Y el número de artículos defectuosos en la muestra y considere un procedimiento que proporciona una región crítica definida por $Y \geq 7$ o $Y \leq 1$.

a) Determine la función de la potencia $\pi(p)$ en los puntos $p \in \{0, 0{,}1, 0{,}2, ..., 0{,}9, 1\}$.
b) Determine el tamaño del test.

6. Sea $x_1, ..., x_n$ una m.a.s. de una distribución normal de media μ desconocida y varianza 1. Sea μ_0 un valor dado. Se tienen las hipótesis

$$H_0 : \mu = \mu_0 \quad vs. \quad H_1 : \mu \neq \mu_0$$

Supongamos que el tamaño de la muestra es 25, y considérese que el procedimiento para no rechazar H_0 está dado por $|\bar{X}_n - \mu_0| < c$. Determine el valor de c para que el tamaño del test sea 0,05.

7. Sea $x_1, ..., x_n$ una m.a.s. de una distribución de media θ desconocida y varianza 1, y sean las hipótesis

$$H_0 : \theta = 3{,}5 \quad vs. \quad H_1 : \theta = 5{,}0$$

a) Entre los procedimientos para resolver el test anterior tal que $\beta(\delta) \leq 0{,}05$, descríbase un procedimiento para el que $\alpha(\delta)$ sea un mínimo.
b) Para $n = 4$, encuéntrese el valor mínimo descrito en a).

8. Supóngase que se selecciona una observación X de una $U(0, \theta)$, donde θ es desconocido y se plantean las siguientes hipótesis:

$$H_0 : \theta = 1 \quad vs \quad H_1 : \theta = 2$$

a) Demuestre que existe un procedimiento para resolver el test para el cual $\alpha(\delta) = 0$ y $\beta(\delta) < 1$.
b) Entre todas las soluciones del test para las cuales $\alpha(\delta) = 0$, hállese una para el cual $\beta(\delta)$ sea mínimo.

9. Sea $x_1, ..., x_n$ una m.a.s. de una $Poisson(\lambda)$, con λ desconocido. Sean λ_0 y λ_1 dados, con $\lambda_1 > \lambda_0 > 0$. Se tienen las siguientes hipótesis:

$$H_0 : \lambda = \lambda_0 \quad vs. \quad H_1 : \lambda = \lambda_1$$

Demuéstrese que el valor de $\alpha(\delta) + \beta(\delta)$ se minimiza por un procedimiento que rechaza H_0 cuando $\bar{X}_n > c$ y encuéntrese el valor de c.

10. Sea $X_1, ..., X_n$ una m.a.s. de una distribución con parámetro θ cuyo valor es desconocido. Supóngase además que se desea contrastar las siguientes hipótesis:

$$H_0 : \theta \leq \theta_0 \quad vs \quad H_1 : \theta > \theta_0$$

Supóngase además, que el procedimiento que se va a utilizar ignora los valores observados en la muestra y, en vez de ello, depende únicamente de una aleatorización auxiliar en la que se lanza una moneda cargada de manera que se obtendrá cara con probabilidad 0.05 y sello con probabilidad 0.95. Si se obtiene una cara, entonces se rechaza H_0, y si se obtiene sello, no se rechaza H_0. Descríbase la función de potencia de este procedimiento.

11. Sea $x_1, ..., x_n$ una m.a.s. de una distribución con parámetro θ desconocido y una función de densidad conjunta $f_n(x|\theta)$ que tiene cociente de verosimilitud monótona en el estadístico $T = r(X)$. Sea θ_0 un valor específico de θ y supóngase que se quieren contrastar las hipótesis

$$H_0 : \theta \geq \theta_0 \quad vs \quad H_1 : \theta < \theta_0$$

Sea c una constante tal que $P(T \leq c|\theta = \theta_0) = \alpha_0$. Demostrar que el procedimiento que rechaza H_0 si $T \leq c$ es U.M.P. al nivel α_0.

12. Sea $x_1, ..., x_n$ una m.a.s. de una $Poisson(\lambda)$ con λ desconocido. Se quiere contrastar las hipótesis

$$H_0 : \lambda \geq 1 \quad vs \quad H_1 : \lambda < 1$$

Supóngase además que el tamaño de la muestra es $n = 20$. ¿Para qué niveles de significación α_0, con $0 < \alpha_0 < 0{,}03$ existen tests UMP?

13. Consideremos una observación X de una distribución de Cauchy con un parámetro de localización desconocido θ, esto es, una distribución cuya función de densidad está dada por:

$$f(x/\theta) = \frac{1}{\pi[1 + (x - \theta)^2]} \quad (\forall x)$$

Se desean contrastar las hipótesis

$$H_0 : \theta = 0 \quad vs \quad H_1 : \theta > 0$$

Demuestre que no existe un test UMP de estas hipótesis a ningún nivel de significación α_0.

14. Sea $x_1, ..., x_n$ una m.a.s. de una distribución $\mathcal{N}(\mu, 1)$. Supóngase que se desean contrastar las hipótesis

$$H_0 : \mu \leq 0 \quad vs \quad H_1 : \mu > 0$$

Se denota δ^* el test UMP con nivel de significación $\alpha_0 = 0{,}025$ y $\pi_{\delta^*}(\mu)$ la función de potencia de δ^*.

15. Sea $x_1, ..., x_n$ una m.a.s. de una distribución $U(0, \theta)$ con θ desconocido. Supongamos que queremos contrastar las hipótesis

$$H_0 : \theta = 3 \quad vs \quad H_1 : \theta \neq 3$$

Considere que H_0 se rechaza si $c_2 \leq max\{x_1, ..., x_n\} \leq c_1$ y sea $\pi(\theta|\delta)$ la función de potencia de δ. Determine los valores de c_1, c_2 para que $\pi(3|\delta) = 0{,}05$ y δ sea insesgado.