Clase 3:
Estadística: Muestras, Intervalos de Confianza y Test de Hipótesis

Pablo Barceló
DCC, Universidad de Chile

1. Estadística Básica de Muestras

Muestras: Una *muestra* es un conjunto $x_1, \ldots, x_n$ de *observaciones*. Como no conocemos a priori el valor de $x_1, \ldots, x_n$ podemos pensar que estas corresponden a los valores de las variables aleatorias $X_1, \ldots, X_n$. Para que $X_1, \ldots, X_n$ sea una *muestra aleatoria* se debe cumplir que los X_i's sean independientes y sean obtenidos con respecto a la misma función de probabilidad.

La media de una muestra: En general, nos interesa la *media* $\bar{X} = (X_1 + \cdots + X_n)/n$ de una muestra.

Importante: Existen relaciones entre el valor esperado y la varianza de la media $\bar{X}$ de una muestra aleatoria $X_1, \ldots, X_n$ obtenida desde una función de probabilidad $f(x)$ con valor esperado μ y varianza σ^2. Note que $\bar{X}$ es una variable aleatoria cuyo experimento corresponde al conjunto de muestras de tamaño n desde $f(x)$. En particular:

1. $E(\bar{X}) = \mu$.
2. $V(\bar{X}) = \sigma^2/n$.

Esto sirve para estimar el valor esperado y la varianza de $f(x)$. Note que el valor esperado de $\bar{X}$ coincide con el valor esperado de $f(x)$ pero que la varianza disminuye a medida que aumenta el número de muestras.

Teorema central del límite: No importando como $f(x)$ se distribuye, a medida que el número n de muestras aumenta la distribución de $\bar{X}$ se va pareciendo cada vez más a una distribución normal (centrada en su valor esperado μ y con varianza σ^2/n cada vez menor).

Esto quiere decir que podemos aproximar la distribución de $\bar{X}$ con una distribución normal para n grande (en la práctica, para $n \geq 30$).

Ejercicio: Una organización reporta el número de defectos importantes de cada auto que testea. La distribución $f(x)$ obtenida tiene valor esperado de 3,2 y desviación estándar 2,4. En una muestra de 100 autos, ¿cuán probable es que la media de defectos por cada auto sea mayor a 4? Respuesta: Podemos asumir que $\bar{X}$ distribuye normal con $\mu = 3,2$ y $\sigma^2 = (2,4)^2$. Por tanto, la desviación estándar de $\bar{X}$ será de $(2,4)/10 = 0,24$. Esto quiere decir que la $Pr(\bar{X} > 4)$ es igual a la probabilidad de que en una distribución normal de estas características el valor de $\bar{X}$ sea mayor a 4. Utilizando métodos analíticos se puede demostrar que esta probabilidad es de 0.0004.

Tamaño de la muestra: Queremos calcular el número de muestras necesarias para asegurarnos de que la probabilidad de que la media $\bar{X}$ difiera mucho del valor esperado μ de la distribución sea a lo más $0 \leq p \leq 1$.

Importante (Inigualdad de Chebyshev): Se cumple que $Pr(|\bar{X} - \mu| \geq t) \leq \sigma^2/nt^2$. Esto nos permite calcular el número necesario de muestras para asegurar una cierta probabilidad.

Ejercicio: Suponga que la varianza de una distribución es a lo más 2 y que queremos determinar el valor de μ con una inexactitud de a lo más 1. ¿Cuantas muestras debemos tomar para que la probabilidad de que eso ocurra sea al menos de 0,99? Respuesta: Utilizaremos $\bar{X}$ para estimar μ. Por tanto, queremos que $Pr(|\bar{X} - \mu| \geq 1) \leq 0,001$. Por la inigualdad, basta con que $\sigma^2/nt^2 = 4/n \leq 0,001$. Es decir, nuestra muestra debe ser de tamaño al menos 400.

2. Estimadores Puntuales

Estimadores: Asuma que θ es un parámetro que queremos conocer desde una distribución $f(x)$. Un *estimador* $\hat{\theta}$ de θ es una variable aleatoria definida en el espacio de las muestras.

Ejemplo: La media $\bar{X}$ de una muestra es un estimador del valor esperado μ de una distribución. Por otra parte, se considera que la varianza $S^2 = \sum_{1 \leq i \leq n} (X_i - \bar{X})^2/(n-1)$ de una muestra es un buen estimador de su varianza σ^2.

Estimadores no sesgados: Son los estimadores que cumplen que $E(\hat{\theta}) = \theta$. Por ejemplo, sabemos que $\bar{X}$ es un estimador no sesgado de μ. Si $\hat{\theta}$ es sesgado, entonces su sesgo corresponde a $E(\hat{\theta}) - \theta$.

Importante: La varianza S^2 de una muestra es un estimador no sesgado de σ^2.

Importante: Idealmente elija siempre estimadores no sesgados.

Ejemplo: Asuma que tenemos una distribución uniforme en el intervalo $[0, \theta]$ y que queremos estimar el valor de θ. Considere el estimador $\hat{\theta}$ definido como $\max\{X_1, \ldots, X_n\}$. ¿Es este estimador no sesgado? Respuesta: No, este valor sí es sesgado. De hecho, su valor esperado es $n \cdot \theta/(n+1)$, que es menor que θ. Su sesgo es $-\theta/(n+1)$. Por otro lado, $\hat{\theta} = (n+1)/n \cdot \max\{X_1, \ldots, X_n\}$ es un estimador no sesgado.

Importante: Cuando existen varios estimadores no sesgados es conveniente utilizar el de menor varianza.

Ejemplo: Para estimar el valor de θ también podemos utilizar $\hat{\theta}' = 2\bar{X}$. Este estimador es no sesgado. Por otro lado, la varianza de $\hat{\theta}$ es $\theta^2/12n$ mientras que la varianza de $\hat{\theta}'$ es $\theta^2/3n$. Es más conveniente, por tanto, utilizar a $\hat{\theta}$ como estimador (de hecho, es posible demostrar que este es el estimador de menor varianza).

Importante: Si $f(x)$ distribuye de forma normal, entonces el mejor estimador para μ es $\bar{X}$.

3. Intervalos de Confianza

Definición: Asumamos que $f(x)$ distribuye normal. Entonces $\bar{X}$ distribuye normal con valor esperado μ y desviación estándar $\sigma/\sqrt{n}$. Esto quiere decir que la variable Z definida como $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ distribuye normal con valor esperado 0 y desviación estándar 1. Para esta distribución es posible conocer los valores de la probabilidad de que $-t \leq Z \leq t$. Por ejemplo, se tiene que $Pr(-1,96 \leq Z \leq 1,96)$ es 0,95. Esto implica que:

$$Pr(-1,96 \cdot \frac{\sigma}{\sqrt{n}} \leq \bar{X} - \mu \leq 1,96 \cdot \frac{\sigma}{\sqrt{n}}) \ = \ 0{,}95.$$

Es decir, $Pr(\bar{X} - 1,96 \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 1,96 \cdot \frac{\sigma}{\sqrt{n}}) \ = \ 0{,}95$.

En general, para cada $0 \leq \alpha \leq 1$ es posible calcular el valor crítico z_α tal que $Pr(Z \geq z_\alpha) = \alpha$. Por tanto, $Pr(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \alpha$. En tal caso tenemos que $Pr(\bar{X} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$. Además, $[\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}]$ se denomina un *intervalo de confianza al* $100(1-\alpha)\,\%$ para μ, donde $\bar{x}$ es la media de alguna muestra.

Interpretación: Esto significa que de todas las muestras de n elementos que saquemos, en el $100(1-\alpha)\,\%$ de ellas tendremos que el valor esperado μ está efectivamente en el intervalo de confianza determinado por la muestra. Es decir, si α es pequeño entonces para una muestra cualquiera existe una buena posibilidad de que el intervalo de confianza calculado contenga a μ.

Ejercicio: Un producto se produce con una desviación estándar igual a $\sigma = 0{,}1$ para cierto parámetro. Al tomar una muestra de 40 casas se obtiene un valor promedio $\bar{x} = 5{,}426$ para tal parámetro. Determine un intervalo de confianza al 90 % para este. Respuesta: En tal caso tenemos que $\alpha = 0{,}1$. Podemos calcular entonces $z_{\alpha/2} = z_{0{,}05} = 1{,}645$. Por tanto, el intervalo de confianza es $5{,}426 \pm 1{,}645 \cdot \frac{0{,}1}{\sqrt{40}} = [5,4; 5{,}452]$.

Importante: Un aumento en confianza significa una pérdida de precisión, a menos que se aumente el número de muestras.

Distribuciones no normales: En el caso de que $f(x)$ no sea normal entonces $\bar{X}$ es “aproximadamente” una distribución normal. Por tanto, el valor del intervalo de confianza es aproximado. Es decir, $Pr(-z_{\alpha/2} \leq Z \leq z_{\alpha_2}) \approx 1 - \alpha$.

Aproximación de la varianza: Cuando σ es desconocido podemos aproximarlo con la desviación S de una muestra. En tal caso agregamos mayor dispersión. Sin embargo, a medida que n crece (por ejemplo, para $n > 40$) la distribución de $\frac{\bar{X}-\mu}{S/\sqrt{n}}$ va aproximándose cada vez más a una distribución normal con media 0 y desviación estándar 1. En particular, $\bar{x} \pm z_{\alpha/2} \cdot S/\sqrt{n}$ es un intervalo de confianza a aproximadamente un $100(1-\alpha)\,\%$ para μ.

4. Test de Hipótesis

Hipótesis: Aseveración acerca del valor de un parámetro. La *hipótesis nula* H_0 es la que se cree cierta inicialmente. La *hipótesis alternativa* H_a es la que contradice H_0. Si nuestra muestra no contradice a H_0 fuertemente, entonces seguiremos creyéndola.

Test de hipótesis: Método para determinar si H_0 debe ser rechazada.

Procedimiento de testeo: Consiste de una *estadística de testeo*, que es una función de la muestra que permite decidir si H_0 será rechazada, y una *región de rechazo*, que es el conjunto de los valores de la estadística que implicarán el rechazo de H_0.

Ejemplo: Un productor de cigarrillos asevera que el procentaje de nicotina por cigarrillo es de 1.5mg. La hipótesis nula, por tanto, es $H_0 : \mu \leq 1{,}5$. La estadística de testeo tomará 32 cigarrillos y medirá el promedio $\bar{x}$ de nicotina en ellos. La región de rechazo será $\bar{x} \geq 1,6$.

Error de testeo: Quizá nuestra muestra es mala y hemos cometido un error al rechazar H_0. Pero, ¿cuán probable es esto? Existen dos tipos de errores:

1. **Tipo I:** Rechazar H_0 siendo esta cierta.
2. **Tipo II:** No rechazar H_0 siendo esta falsa.

La probabilidad de error tipo I y II se denota por α y β, respectivamente.

Ejercicio: El 25 % de los autos de un tipo que tienen un accidente a 20 kms/h no presenta daño visible. Se diseña un nuevo tipo de protección que asevera aumentar considerablemente este porcentaje. Se probarán 20 autos con la protección y se aceptará esta como confiable si al menos 8 de ellos no presentan daño visible (es decir, al menos el 40 %). Calcule los valores de α y β, respectivamente. Respuesta: La hipótesis H_0 es $p = 0{,}25$. Sea X el número de automóviles testeados que no presentaron daño. Entonces, $\alpha = Pr(H_0$ se rechaza pero es verdadera) es igual a la probabilidad de que $X \geq 8$ dado que la probabilidad de no tener daño es a lo más 25 %. Esto puede calcularse como 0.102. Por otro lado, no existe un único valor para β sino que este depende del valor $> 0{,}25$ que tome p. Por ejemplo, para $p = 0{,}3$ tenemos que $\beta = Pr(H_0$ no se rechaza pero la hipótesis es falsa debido a que $p = 0{,}3)$. Esto es igual a la probabilidad de que $X \leq 7$ dado que la probabilidad de no tener daño es mayor a 0.3. Esta probabilidad es $0{,}772$. Cuando $p = 0{,}4$ esta probabilidad disminuye a $0{,}416$.

Ejercicio: Suponga que el tiempo de secado de un tipo de pintura tiene valor esperado de 75 min y desviación estándar de 9 minutos. Se propone un aditivo para mejorar el tiempo promedio μ de secado. Se asume que la distribución con el aditivo sigue siendo normal con $\sigma = 9$. Para testear el aditivo se harán 25 pruebas y se calculará el promedio $\bar{X}$ de tiempo de secado. La hipótesis $\mu = 75$ se rechazará solo si $\bar{X} \leq 70{,}8$. Calcule valores para α y β. Respuesta: Sabemos que $\bar{X}$ distribuye normal con valor esperado μ y desviación estándar $9/\sqrt{25} = 1{,}8$. Para α computamos la probabilidad de que $\bar{X} \leq 70{,}8$ dado que $\mu = 75$. Es decir, queremos saber la probabilidad de que $\bar{X} \leq 70{,}8$ cuando $\bar{X}$ distribuye normal con valor esperado 75 y desviación 1.8. Esto es lo mismo que calcular la probabilidad de que $Z \leq (70{,}8 - 75)/1{,}8$ en una distribución normal con media 0 y desviación 1, lo que puede calcularse como $0{,}01$. Para el valor de β debemos especificar cuál es el valor de μ que lo causa. Por ejemplo, si $\mu = 72$ entonces β corresponde a la probabilidad de que $X \geq 70{,}8$ dado que $\mu = 72$. Esto corresponde a la probabilidad de que $Z \geq \frac{70{,}8-72}{1{,}8}$. Esto puede calcularse como $0{,}7486$. Si disminuye el valor de μ entonces también disminuye este error.

Pregunta: ¿Cuál de los dos errores se considera usualmente más peligroso? Respuesta: El de tipo I. Pero si queremos disminuír el valor de α esto solo puede hacerse al costo de aumentar β. Idea: Encontrar el mayor valor de α aceptable y ajustar el test a ese valor. En tal caso decimos que α es el *nivel de significancia* del test y el test es un *test de nivel* α.

Ejemplo: Sigamos con el ejemplo del porcentaje de nicotina en los cigarros. Suponga que la distribución de nicotina es normal con desviación estándar $\sigma = 0,2$. Por tanto, $\bar{X}$ distribuye normal con la misma media μ que la distribución original y desviación estándar $0,2/\sqrt{32} = 0,0354$. Si H_0 es verdadera, es decir, $\mu = 1,5$, entonces nuestra estadística será $Z = (\bar{X} - 1,5)/0,0354$. Asuma que el máximo error de tipo I que permitimos es $\alpha = 0,05$. Entonces nuestro "valor de decisión" para el test debe ser el valor c tal que $Pr(Z \geq c) = 0,05$, cuando Z distribuye normal con media 0 y desviación estándar 1. Tal valor de c corresponde a $1,645$. Resolviendo para $\bar{X}$ obtenemos que la región de rechazo debería ser $\bar{X} \geq 1,645 \cdot 0,0354 + 1,5$, es decir, aproximadamente $1,56$.

5. Procedimientos de Testeo

Distribución normal con desviación σ conocida: Hipótesis nula $H_0 : \mu = \mu_0$. Sea $\bar{X}$ la media de una muestra aleatoria $X_1, \ldots, X_n$. Si H_0 es verdadera entonces $\mu_{\bar{X}} = \mu_0$. Además, $\sigma_{\bar{X}} = \sigma/\sqrt{n}$. Definimos entonces la variable

$$Z = \frac{(\bar{X} - \mu)}{\sigma/\sqrt{n}},$$

la que expresa la diferencia entre $\bar{X}$ y μ en unidades de la desviación estándar. Si Z es muy grande (o muy chica) rechazamos H_0. Por tanto, en este caso el test se realiza de la siguiente forma:

Hipótesis nula:	$H_0 : \mu = \mu_0$
Estadística de testeo:	$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$
Hipótesis alternativa	Región de rechazo para test de nivel α
$H_a : \mu > \mu_0$	$z \geq z_\alpha$
$H_a : \mu < \mu_0$	$z \leq -z_\alpha$
$H_a : \mu \neq \mu_0$	$z \leq -z_{\alpha/2}$ o $z > z_{\alpha/2}$

El procedimiento se realiza de la siguiente forma:

1. Identificar el parámetro de interés.
2. Determinar el valor nulo y estipular la hipótesis nula.
3. Estipular la hipótesis alternativa adecuada.
4. Dar una fórmula para el valor de la estadística a computar.
5. Establecer la región de rechazo para el nivel de significancia α.
6. Computar los valores necesarios de la muestra y reemplazarlos en la fórmula de la estadística para obtener un valor.
7. Decidir si la hipótesis nula debe ser rechazada.

Importante: Los pasos (2) y (3) deben ser realizados sin mirar a los datos.

Ejemplo: Se dice que la temperatura promedio de activación para un sistema anti-incendios es de 130°F. Una muestra de 9 sistemas muestra un promedio de activación de 131.08°F. Si la distribución de los tiempos de activación es normal con $\sigma = 1{,}5°F$, ¿existe evidencia contradictoria para la aseveración inicial con $\alpha = 0{,}01$?

1. Parámetro de interés: μ.

2. Hipótesis nula: $H_0 : \mu = 130$.
3. Hipótesis alternativa: $H_a : \mu \neq 130$.
4. Estadística de testeo: $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{\bar{x}-130}{1{,}5/\sqrt{n}}$.
5. Región de rechazo: $z_{0{,}005} \leq z$ o $z \leq -z_{0{,}005}$. Es decir, $z \geq 2,58$ o $z \leq -2,58$.
6. Reemplazar $\bar{x} = 131{,}08$ y $n = 9$ en la estadística de testeo. Esto da $z = 2,16$.
7. No rechazamos ya que z no cae en región de rechazo.

Importante: El valor de α ya está dado. ¿Cómo podemos determinar el valor de β. Asuma que tenemos un test con región de rechazo $z \geq z_\alpha$. Esto es equivalente a $\bar{x} \geq \mu_0 + z_\alpha \cdot \sigma/\sqrt{n}$. Es decir, H_0 no será rechazada si $\bar{x} < \mu_0 + z_\alpha \cdot \sigma/\sqrt{n}$. Determinemos $\beta(\mu')$, para $\mu' > \mu_0$. Entonces:

$$\begin{aligned}
\beta(\mu') &= Pr(\text{no rechazar } H_0 \text{ cuando } \mu = \mu') \\
&= Pr(\bar{X} < \mu_0 + z_\alpha \cdot \sigma/\sqrt{n} \text{ cuando } \mu = \mu') \\
&= Pr(\frac{\bar{X}-\mu'}{\sigma/\sqrt{n}} < \frac{\mu_0-\mu'}{\sigma/\sqrt{n}} + z_\alpha)) \\
&= \Phi(\frac{\mu_0-\mu'}{\sigma/\sqrt{n}} + z_\alpha).
\end{aligned}$$

Para los otros tests el valor se encuentra en la siguiente tabla:

Hipótesis alternativa	Error de tipo II $\beta(\mu')$ para el test de nivel α
$H_a : \mu > \mu_0$	$\Phi(\frac{\mu_0-\mu'}{\sigma/\sqrt{n}} + z_\alpha)$
$H_a : \mu < \mu_0$	$1 - \Phi(\frac{\mu_0-\mu'}{\sigma/\sqrt{n}} - z_\alpha)$
$H_a : \mu \neq \mu_0$	$\Phi(\frac{\mu_0-\mu'}{\sigma/\sqrt{n}} + z_{\alpha/2}) - \Phi(\frac{\mu_0-\mu'}{\sigma/\sqrt{n}} - z_{\alpha/2})$

Ejercicio: Sea μ el tiempo esperado de duración de un cierto tipo de neumático. Considere testear $H_0 : \mu = 30000$ vs $H_a : \mu > 30000$ desde una muestra de tamaño $n = 16$. Asumimos que la distribución es normal con $\sigma = 16$. Queremos hacer un test de nivel $\alpha = 0{,}01$. ¿Cuál es el valor de $\beta(31000)$? ¿Cuántas muestras son necesarias para asegurarnos de que β tenga valor 0,1? Respuesta: $\beta(31000) = \Phi(2,33 + \frac{30000-31000}{1500/\sqrt{16}}) = \Phi(-0{,}34) = 0{,}3669$. Si ahora queremos que $\Phi(2,33 + \frac{30000-31000}{1500/\sqrt{n}}) = 0{,}1 = \Phi(-1{,}28)$ se puede concluir que n debe valer al menos 30.

Distribución no normal y/o varianza desconocida: En el primer caso, si el número de muestras es grande, podemos aproximar por una normal. En el segundo caso utilizamos la estimación S de la desviación estándar. Lo único que cambia es el hecho de que el nivel de significancia es ahora *aproximadamente* α.

Valores-p: Otra forma de llegar a una conclusión con respecto a la hipótesis H_0. El valor-p es la probabilidad de obtener un valor de la estadística al menos tan contradictorio con H_0 como el valor calculado de nuestra muestra, asumiendo que H_0 es cierta.

Ejemplo: Una muestra de 51 pilas AAA muestra que su masa promedio de zinc es $2{,}06g$ con una desviación en la muestra de $0{,}141g$. Hay suficiente evidencia en estos datos para concluír que el valor medio de la cantidad de zinc en una pila es mayor a 2? Utilizaremos $H_0 : \mu = 2$ y $H_a : \mu > 2$. Dado que la muestra es suficientemente grande podemos asumir que la estadística $z = \frac{\bar{x}-\mu}{s/\sqrt{n}}$ pueda ser utilizada sin asumir nada sobre la distribución. Note que en este caso $z = \frac{2{,}06-2{,}0}{0{,}141/\sqrt{51}} = 3{,}04$. Note que cualquier valor de $\bar{x} \geq 2{,}06$ es al menos tan contradictorio como $2{,}06$ a H_0. Pero cualquiera de esos valores corresponde a un valor de $z \geq 3{,}04$, es decir, a un valor de la estadística que es al menos tan contradictorio con H_0 como $3{,}04$, asumiendo que $\mu = 2$. Por tanto, el valor-p en este caso es

$Pr(Z \geq 3{,}04$ dado que $\mu = 2)$. Como Z tiene aproximadamente una distribución normal estándar, tenemos que esto último corresponde a $1 - \Phi(3{,}04) = 0{,}0012$.

Pregunta: ¿Qué nos dice este, o cualquier otro valor-p, sobre cuánto podemos confiar en H_0? Considere $p = 0{,}25$. Entonces un 25 % de los valores de la estadística son más contradictorios con H_0 que el valor computado desde nuestra muestra. Es decir, nuestro valor no parece tan contradictorio con H_0. Por otro lado, si $p = 0{,}0018$ entonces nuestro valor calculado es muy contradictorio con H_0.

Importante: Mientras más pequeño es el valor-p más evidencia entrega nuestra muestra en contra de H_0.

Regla de decisión basada en valor-p:

1. Seleccione un nivel de significancia α
2. Rechace si: valor-$p \leq \alpha$
3. No rechace si: valor-$p > \alpha$

Importante: Note que $\alpha < p$ si y solo si $z_p < z_\alpha$. Por tanto, el test usual rechaza H_0 si y solo si el test basado en el valor p rechaza H_0. Además, p es el menor nivel de significancia α al que H_0 puede ser rechazada. Comúnmente por esto llamamos al valor-p el *nivel observado de significancia*.

Ejercicio: El valor esperado que tarda en hacer efecto una aspirina es 10 minutos. Sea μ el valor esperado de este valor para un nuevo medicamento. Entonces $H_0 : \mu = 10$ y $H_a : \mu < 10$. El valor-p calculado para esta hipótesis es de $0{,}0384$. Es decir, solo un 3 % de los valores calculados de la estadística son más contradictorios con H_0 que el valor calculado desde la muestra. Considere $\alpha = 0{,}05$. Dado que $p \leq \alpha$, el test basado en el valor-p rechaza la hipótesis. De igual forma lo haríamos si realizaramos el test usual con nivel de significancia α. Por otro lado, si $\alpha = 0{,}01$ ninguno de los tests rechaza la hipótesis.

Importante: Los valores-p permiten tomar decisiones solo en referencia a α, sin hacer referencia a los valores críticos. Además, permiten distinguir decisiones muy claras de otras más difusas en términos del rechazo o aceptación de la hipótesis.

Valores-p para tests z: Este es el caso en que la estadística cuando H_0 es verdadera es aproximadamente normal. En general:

$$\boxed{\text{valor-}p \ = \begin{cases} 1 - \Phi(z), & \text{si } H_a : \mu \geq \mu_0 \\ \Phi(z), & \text{si } H_a : \mu \leq \mu_0 \\ 2(1 - \Phi(|z|)), & \text{si } H_a : \mu \neq \mu_0 \end{cases}}$$

Ejercicio: Un fabricante de silicona afirma que el grosor medio de sus cables es de 245um. Una muestra de 50 cables entrega un valor medio de 246.18uM y una desviación en la muestra de 3.6uM. ¿Es esta muestra sugerente del hecho que el valor promedio del grosor es distinto de 245uM? Respuesta:

1. Parámetro de interés: el grosor medio μ de un cable.
2. Hipótesis nula: $H_0 : \mu = 245$.
3. Hipótesis alternativa: $H_a : \mu \neq 245$.
4. Estadística: $z = \frac{\bar{x} - 245}{s/\sqrt{n}}$.
5. Valor de la estadística en la muestra: $z = \frac{246{,}18 - 245}{3{,}6/\sqrt{50}} = 2{,}32$.
6. Valor-p: $2(1 - \Phi(2{,}32)) = 0, 0204$.
7. Conclusión: El menor nivel de significancia α al cual H_0 sería rechazado es $0{,}0204$.