Resumen Tests de Hipótesis

18 de diciembre de 2016

Acerca del p-valor:

El p-valor puede ser descrito como la probabilidad, bajo la hipótesis nula (H_o), de obtener resultados tan extremos como los observados. Lo que buscamos con el test de hipótesis es determinar esta probabilidad, y si es muy baja (menos que el nivel de confianza del test, α) rechazar la hipótesis nula. La lógica detrás es que si esta probabilidad es muy baja, significa que es muy difícil haber obtenido los datos que tenemos si H_o fuera cierta, por lo tanto es probable que sea falsa (es como la versión probabilística de la contradicción: si P $\implies$ Q y Q es falso, entonces P es falso. Solo que en este caso no tenemos los valores de verdad, sino que una probabilidad de ellos).

El nivel de significancia (α) es la probabilidad de cometer error tipo 1 (rechazar la hipótesis nula cuando esta es verdadera) y es un valor determinado por quien realiza el experimento, generalmente antes de realizarlo, y tiene que ver con cuán frecuente se quiere que el test no se equivoque del tipo 1. Por ejemplo, un test de VIH con $\alpha = 0{,}05$, dice que el paciente tiene VIH cuando no lo tiene (H_o), un 5 % de las veces.

En un test de hipótesis, si el p-valor es menor que el nivel de significancia deseado, se rechaza la hipótesis nula. Ya que la probabilidad de rechazar la hipótesis nula cuando esta es cierta es menor que el nivel deseado de esta probabilidad. Por esto, el p-valor, también puede ser visto como el menor nivel de significancia del test para el cuál se rechaza la hipótesis nula. Ejemplo: Si el p-valor resulta 0.001, entonces el test se rechazará para todos los niveles de significancia superiores a 0.001, como 0.01, 0.05, 0.005, etc. Por esto es que el nivel de significancia suele decidirse antes de realizar el experimento, para evitar que el investigador lo decida en base a los p-valores obtenidos y las que hipótesis le gustaría rechazar (si el p-valor fuera 0.03, un test al 5 % rechazaría H_o, mientras que un test al 2.5 % no).

Efectos prácticos:

Supongamos z un estadístico creciente con respecto al estimador del parámetro a testear (por ejemplo, $z = \frac{(n-1)S^2}{\sigma^2}$ es creciente con respecto a S^2, el estimador de σ^2, la varianza de una normal). Si el parámetro es θ, entonces para el test de hipótesis:

$$H_o\colon \theta = \theta_o$$
$$H_1\colon \theta > \theta_o$$

El p-valor estaría dado por $\mathbb{P}(z > z_{obs}|H_o)$

Donde z es el estadístico visto como variable aleatoria, y z_{obs} es el estadístico evaluado en los datos observados. Así es claro ver la primera descripción del p-valor: la probabilidad de observar datos, "$\mathbb{P}(z$", tan extremos, " $>$ ", como los datos observados, "z_{obs}", si la hipótesis nula fuera cierta, "$|H_o)$".

Para el caso en que H_1 es la otra desigualdad, se ocupa $\mathbb{P}(z < z_{obs}|H_o)$. Si H_1 fuera " $\neq$ ", entonces se realiza el test para " $>$ " y para " $<$ ", pero con la mitad de la significancia inicial en cada uno.

Resumen

Test	Datos necesarios	Parámetro a testear	H$_o$	Estadístico	Distr. Estadístico bajo H$_o$
(1)	$\bar{X}, \sigma^2, X \sim N(\mu, \sigma^2)$	μ	$\mu = \mu_o$	$\frac{\bar{X}-\mu_o}{\sigma/\sqrt{n}}$	$N(0,1)$
(2)	$\bar{X}, S^2, X \sim N(\mu, \sigma^2)$	μ	$\mu = \mu_o$	$\frac{\bar{X}-\mu_o}{s/\sqrt{n}}$	t_{n-1}
(3)	$\bar{X}, S^2, X \sim N(\mu, \sigma^2)$	σ^2	$\sigma^2 = \sigma_o^2$	$\frac{(n-1)S^2}{\sigma_o^2}$	χ^2_{n-1}
(4)	$\hat{p}$	p	$p = p_o$	$\frac{\hat{p}-p_o}{\sqrt{p(1-p/n}}$	$N(0,1)$
(5)	$\hat{p_j}, \forall j \in \{1..m\}$	p_j	$\forall i \in \{1..m\} p_i = p$	$n\sum_{j=1}^{m}\frac{(\hat{p_j}-p)^2}{p}$	χ^2_{m-1}
(6)	$\bar{X}, \bar{Y}, S_{x+y}$	μ_x, μ_y	$\mu_x = \mu_y$	$\pm\frac{\bar{X}-\bar{Y}}{S_{x+y}\sqrt{\frac{1}{n_x}+\frac{1}{n_y}}}$	t_{n+m-2}

Explicación:
(1) Test de hipótesis para la media de una Normal con varianza conocida.
(2) Test de hipótesis para la media de una Normal con varianza desconocida.
(3) Test de hipótesis para la varianza de una Normal.
(4) Test de hipótesis para una Bernoulli(p).
(5) Test de bondad de ajuste. m representa el número de alternativas, n el número de muestras total, $\hat{p_j}$, es el estimador de la probabilidad de elegir la alternativa j. Acá se testea si todas las probabilidades son iguales, o si hay alguna distinta. En este test, el p-valor siempre está dado por $\mathbb{P}(z > z_{obs}|H_o)$ (notar el sentido de la desigualdad), ya que bajo la hipótesis nula ($p_j \mathrel{\hat{=}} p, \forall j$), el estadístico debería resultar cercano a 0. Por otro lado, si alguno es distinto, los términos al cuadrado crecen y el estadístico solo puede aumentar.
(6) Test de igualdad de medias. Este test pretende verificar si las medias de dos distribuciones son iguales. S_{x+y} es la desviación estándar de las muestras unidas, se calcula como $\frac{(n_x-1)S_x+(n_y-1)S_y}{n_x+n_y-2}$, donde S_x y S_y son

las desviaciones de las muestras por separado y n_x y n_y el número de datos de cada una (los $(n-1)$ permiten que S_{x+y} sea un estimador insesgado de la desv. estándar de la población de los X e Y unidos). Cuando H_1 es $\mu_x > \mu_y$ se ocupa el estadístico $\frac{\bar{X}-\bar{Y}}{S_{x+y}\sqrt{\frac{1}{n_x}+\frac{1}{n_y}}}$ y cuando H_1 es $\mu_x < \mu_y$ se ocupa $\frac{\bar{Y}-\bar{X}}{S_{x+y}\sqrt{\frac{1}{n_x}+\frac{1}{n_y}}}$ (notar la simetría entre X e Y).

Nota: Para el examen, la hipótesis alternativa estará dada por el enunciado. Por ejemplo, si la discusión es acerca de que un parámetro es mayor que otro, menor que otro, distinto de otro o si hay razones para asumir que la diferencia solo puede ir en un sentido.