Docimasía de Hipótesis

Hipótesis estadística es una afirmación respecto de una característica poblacional (forma de ella o valor de sus parámetros); esta sentencia puede ser "docimada" (probada) en base a una muestra aleatoria extraída de esa población.

En muchas ocasiones es necesario decidir entre una afirmación de la forma $\theta = \theta_0$ (Hipótesis nula) u otra que puede tomar las siguientes formas $\theta \neq \theta_0, \theta > \theta_0, \theta < \theta_0$ (Hipótesis alternativa). En símbolos:

$$
\begin{aligned}
&H_0 : \theta = \theta_0 \\
&H_1 : \theta \neq \theta_0 \\
&ó \\
&H_1 : \theta > \theta_0 \\
&ó \\
&H_1 : \theta < \theta_0
\end{aligned}
$$

Necesitamos desarrollar un procedimiento que nos permita tomar una decisión acerca de H_0, como esta decisión es tomada en base a información muestral está sujeta a errores probables, debido a que no se sabe cómo es realmente la naturaleza pues sólo tenemos una percepción de ella, cruzando este efecto con la decisión tenemos:

		Estado de la naturaleza	
		H_o es Verdad	H_o es Falsa
Percepción de la naturaleza	Rechazar H_o	Error tipo I	Decisión correcta
	No rechazar H_o	Decisión correcta	Error tipo II

Deseamos que los errores no se cometan, pero como la decisión será tomada bajo incertidumbre, sólo podemos pedir que la probabilidad de cometerlos sea pequeña.

La filosofía para docimar consiste en suponer que H_o es verdadera, hasta encontrar evidencia muestral suficiente que permita decir lo contrario, si esta evidencia no existe no podemos dudar de la afirmación contenida en H_o. Así el error más grave que se puede cometer es el Error tipo I, que es el que tratamos de controlar. Llamando:

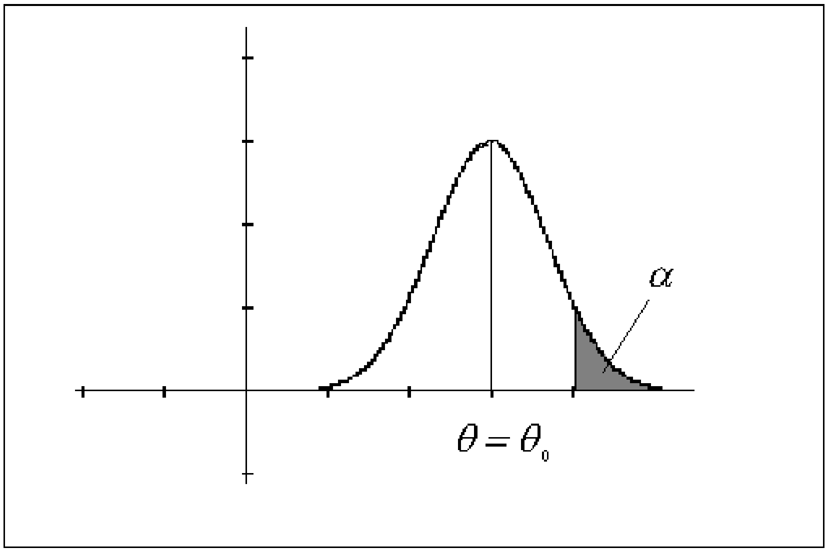
α=P(Rechazar H_o | H_o es Verdad) , tamaño del Error tipo I

β=P(No rechazar H_o | H_o es Falsa) , tamaño del Error tipo II

nos interesa que α sea pequeño α se llama significación de la dócima y 1-β se llama potencia de la dócima, la potencia depende de la hipótesis alternativa que estemos proponiendo y puede interpretarse como el grado de credibilidad que le asignamos a la hipótesis alternativa.

Se llama estadística de prueba, E, a una función que contenga el parámetro de interés (que se desea docimar) y toda la información muestral. Deseablemente la estadística de prueba, bajo la hipótesis nula, debe seguir una distribución de probabilidades conocida.

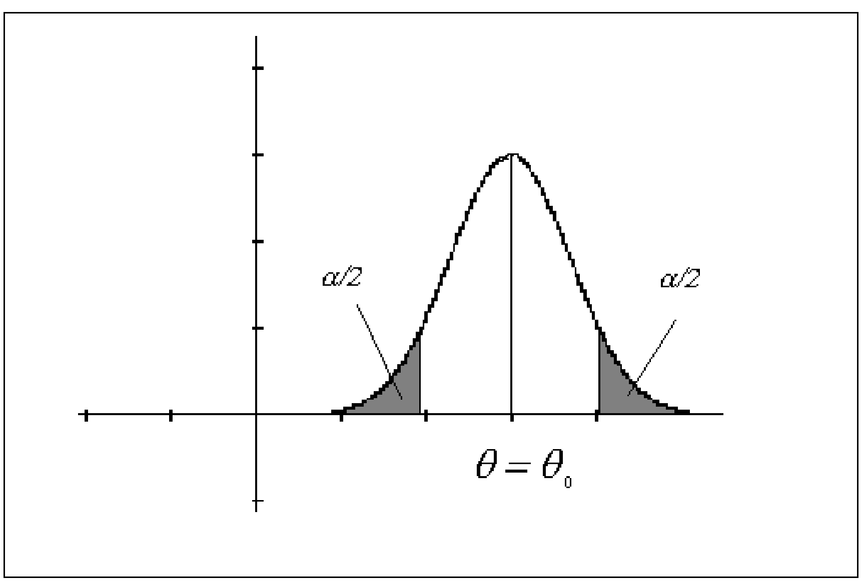
Se llama región crítica o de rechazo, aquella porción de los reales para la cual la probabilidad de que E esté en ella, considerando la veracidad de H_0, sea menor que α



- Una dócima de la forma:

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta \neq \theta_0$$

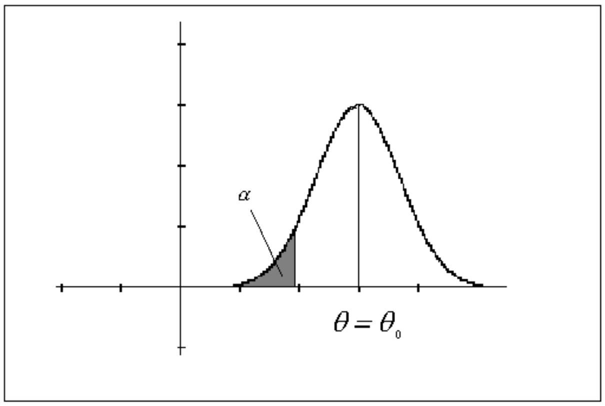
se llama de "dos colas" pues la región de rechazo, se compone de dos porciones de los reales inconexas, que se muestran en el siguiente gráfico:

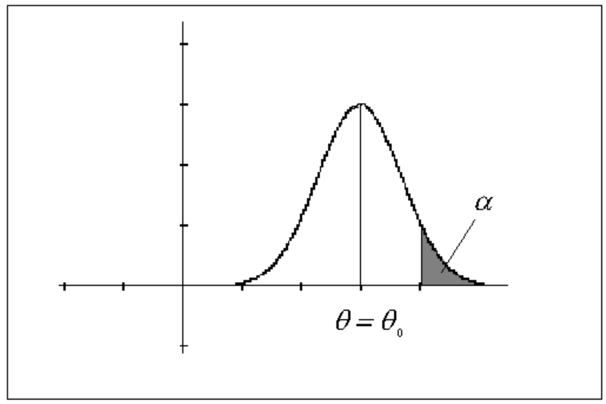


- Una dócima de la forma:

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta > \theta_0$$
$$ó$$
$$H_1 : \theta < \theta_0$$

se llama de "una cola" pues la región de rechazo, se compone de una porción de los reales conexa, como se muestra a continuación:





$H_1 : \theta < \theta_0$ $\qquad\qquad$ $H_1 : \theta > \theta_0$

- Dócimas respecto de promedios:

Hipótesis Nula	Estadística de Prueba	Distribución de la estadística de prueba
$H_0 : \mu = \mu_0$	$\dfrac{\overline{X} - \mu_0}{\dfrac{S}{\sqrt{n}}}$	$t_{(n-1)}$
$H_0 : \mu_x - \mu_y = 0$	$\dfrac{\overline{X} - \overline{Y}}{S_c \sqrt{\dfrac{1}{n_x} + \dfrac{1}{n_y}}}$ $S_c = \sqrt{\dfrac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2}}$	$t_{(n_x + n_y - 2)}$

- Dócimas respecto de varianzas:

Hipótesis Nula	Estadística de Prueba	Distribución de la estadística de prueba
$H_0 : \sigma^2 = \sigma_0^2$	$\dfrac{(n-1)S^2}{\sigma_0^2}$	$\chi^2_{(n-1)}$
$H_0 : \dfrac{\sigma_x^2}{\sigma_y^2} = 1$	$\dfrac{S_x^2}{S_y^2}$	$F_{(n_x - 1, n_y - 1)}$

- Dócimas respecto de proporciones:

Hipótesis Nula	Estadística de Prueba	Distribución de la estadística de prueba
$H_0 : P = P_0$	$\dfrac{p - P_0}{\sqrt{\dfrac{P_0 Q_0}{n}}}$	$N(0,1)$
$H_0 : P_x - P_y = 0$	$\dfrac{p_x - p_y}{\sqrt{PQ(\dfrac{1}{n_x} + \dfrac{1}{n_y})}}$ $P = \dfrac{n_x p_x + n_y p_y}{n_x + n_y}$	$N(0,1)$

Las dócimas anteriores en STATA:

Usando la base de datos stress_cortisol.dta, que contiene la siguiente información:

```
  obs:           500
 vars:             7                          22 Apr 2009 17:48
 size:        16,000 (99.9% of memory free)
-------------------------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------------------------
id              float  %9.0g                  identificador
region          float  %9.0g                  región
sexo            float  %9.0g                  0: masc 1: fem
edad            float  %9.0g                  años cumplidos
stress          float  %9.0g                  0: sin stress 1: con stress
cortisol        float  %9.0g                  nivel de cortisol en mmg
exam            float  %9.0g                  0:neg 1:pos
--------------------------------------------------------------------------------
```

Docimar:

a) $H_0: \mu_{edad} = 30 \ años$

```
. ttest edad=30

One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
    edad |     500       49.55    .3551158     7.94063    48.85229    50.24771
------------------------------------------------------------------------------
    mean = mean(edad)                                             t =  55.0525
Ho: mean = 30                                    degrees of freedom =      499

    Ha: mean < 30            Ha: mean != 30              Ha: mean > 30
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

b) $H_0: \mu_{\text{edad en hombres}} = \mu_{\text{edad en mujeres}}$

```
. ttest edad, by(sexo)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     240     49.2875    .4923962    7.628169    48.31751    50.25749
       1 |     260    49.79231    .5101405    8.225768    48.78776    50.79686
---------+--------------------------------------------------------------------
combined |     500       49.55    .3551158     7.94063    48.85229    50.24771
---------+--------------------------------------------------------------------
    diff |           -.5048077     .711154                -1.90204    .8924244
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -0.7098
Ho: diff = 0                                     degrees of freedom =      498

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.2391         Pr(|T| > |t|) = 0.4781          Pr(T > t) = 0.7609
```

c) $H_0: P_{\text{proporción de stress}} = 0.5$

```
. prtest  stress=.5

One-sample test of proportion                 stress: Number of obs =      500
------------------------------------------------------------------------------
    Variable |       Mean   Std. Err.                     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      stress |       .092   .0129256                      .0666662    .1173338
------------------------------------------------------------------------------
    p = proportion(stress)                                    z = -18.2463
Ho: p = 0.5

     Ha: p < 0.5                 Ha: p != 0.5                 Ha: p > 0.5
 Pr(Z < z) = 0.0000         Pr(|Z| > |z|) = 0.0000          Pr(Z > z) = 1.0000
```

d) $H_0: P_{\text{proporción de stress en hombres}} = P_{\text{proporción de stress en mujeres}}$

```
. prtest  stress, by( sexo)

Two-sample test of proportion                      0: Number of obs =      240
                                                   1: Number of obs =      260
------------------------------------------------------------------------------
    Variable |       Mean   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           0 |   .0791667   .0174283                      .0450077    .1133256
           1 |   .1038462   .0189191                      .0667655    .1409269
-------------+----------------------------------------------------------------
        diff |  -.0246795   .0257231                     -.0750959    .0257369
             |  under Ho:    .025872    -0.95   0.340
------------------------------------------------------------------------------
        diff = prop(0) - prop(1)                                  z =  -0.9539
    Ho: diff = 0

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(Z < z) = 0.1701         Pr(|Z| < |z|) = 0.3401          Pr(Z > z) = 0.8299
```

e) $H_0: \frac{\sigma^2_{\text{edad en hombres}}}{\sigma^2_{\text{edad en mujeres}}} = 1$

```
. sdtest  edad, by(sexo)

Variance ratio test
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     240     49.2875    .4923962    7.628169    48.31751    50.25749
       1 |     260    49.79231    .5101405    8.225768    48.78776    50.79686
---------+--------------------------------------------------------------------
combined |     500       49.55    .3551158     7.94063    48.85229    50.24771
------------------------------------------------------------------------------
    ratio = sd(0) / sd(1)                                         f =   0.8600
Ho: ratio = 1                                   degrees of freedom = 239, 259

    Ha: ratio < 1              Ha: ratio != 1                 Ha: ratio > 1
  Pr(F < f) = 0.1181         2*Pr(F < f) = 0.2362           Pr(F > f) = 0.8819
```

Nota importantísima: recuerde que si las varianzas en dos grupos, entre los cuales se desea comparar promedios, es significativamente distinta, el test t-Student hay que corregirlo, como se muestra en la siguiente colección de datos:

```
     +-----------------+
     | id   grupo    x |
     |-----------------|
  1. |  1       0   19 |
  2. |  2       0   34 |
  3. |  3       0   31 |
  4. |  4       0   33 |
  5. |  5       0   35 |
     |-----------------|
  6. |  6       0   18 |
  7. |  7       0   33 |
  8. |  8       0   15 |
  9. |  9       0   31 |
 10. | 10       0   41 |
     |-----------------|
 11. | 11       1   38 |
 12. | 12       1   26 |
 13. | 13       1   45 |
 14. | 14       1   62 |
 15. | 15       1   73 |
     |-----------------|
 16. | 16       1   48 |
 17. | 17       1   46 |
 18. | 18       1   75 |
 19. | 19       1   19 |
 20. | 20       1   59 |
     +-----------------+


. sort  grupo

. by grupo: sum x

---------------------------------------------------------------------------------------
-> grupo = 0

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
           x |        10          29    8.576454         15         41

---------------------------------------------------------------------------------------
-> grupo = 1

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
           x |        10        49.1    18.54993         19         75

. sdtest  x, by(grupo)

Variance ratio test
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |      10          29    2.712113    8.576454    22.86477    35.13523
       1 |      10        49.1    5.866004    18.54993    35.83018    62.36982
---------+--------------------------------------------------------------------
combined |      20       39.05    3.899713    17.44005    30.88781    47.21219
------------------------------------------------------------------------------
    ratio = sd(0) / sd(1)                                         f =   0.2138
Ho: ratio = 1                                    degrees of freedom =     9, 9

    Ha: ratio < 1               Ha: ratio != 1                 Ha: ratio > 1
  Pr(F < f) = 0.0156         2*Pr(F < f) = 0.0312          Pr(F > f) = 0.9844
```

El p-value en negritas indica que hay evidencia para creer que las varianzas en los grupos es distinta, en estas condiciones el test t-Student es:

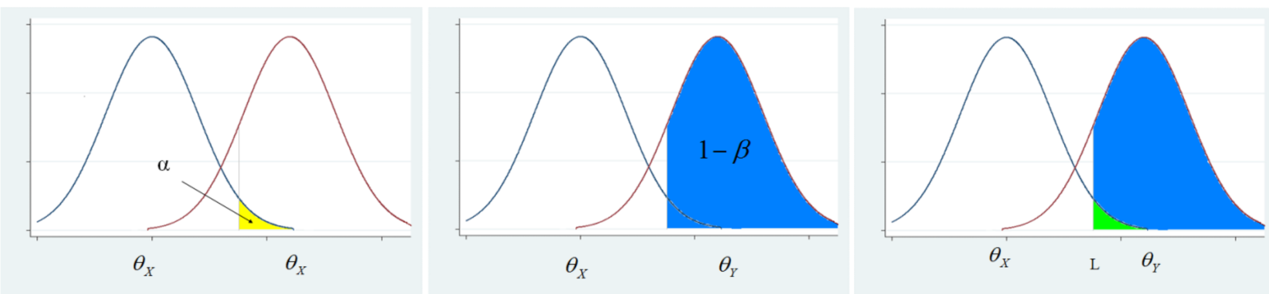
```
. ttest  x, by(grupo) unequal

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |      10          29    2.712113    8.576454    22.86477    35.13523
       1 |      10        49.1    5.866004    18.54993    35.83018    62.36982
---------+--------------------------------------------------------------------
combined |      20       39.05    3.899713    17.44005    30.88781    47.21219
---------+--------------------------------------------------------------------
    diff |               -20.1    6.462628               -34.09761   -6.102392
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -3.1102
Ho: diff = 0                     Satterthwaite's degrees of freedom =  12.6796

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0043         Pr(|T| > |t|) = 0.0085          Pr(T > t) = 0.9957

. ttest  x, by(grupo) unequal welch

Two-sample t test with unequal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |      10          29    2.712113    8.576454    22.86477    35.13523
       1 |      10        49.1    5.866004    18.54993    35.83018    62.36982
---------+--------------------------------------------------------------------
combined |      20       39.05    3.899713    17.44005    30.88781    47.21219
---------+--------------------------------------------------------------------
    diff |               -20.1    6.462628               -34.00955    -6.19045
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -3.1102
Ho: diff = 0                             Welch's degrees of freedom =  13.4973

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0040         Pr(|T| > |t|) = 0.0080          Pr(T > t) = 0.9960
```

El tamaño de la muestra para detectar diferencias de promedios o proporciones en muestras independientes

Como ya hemos visto, si $\hat{\theta}$ es el estimador del parámetro θ, la varianza del estimador involucra al tamaño de la muestra "n", para el caso del promedio muestral su varianza es $\frac{\sigma^2}{n}$, y, para el caso de la proporción muestral su varianza es $\frac{p(1-p)}{n}$. Con estas consideraciones supongamos que se desea encontrar el número de sujetos necesarios a reclutar en dos grupos para hacer la siguiente dócima:

$$H_0: \theta_X - \theta_y = 0$$
$$H_1: \theta_X - \theta_y = \delta$$

Es decir deseamos detectar una diferencia de "δ", con una significación "α" y una potencia "1-β". Sin pérdida de generalidad, supongamos que δ<0 (es decir θ en la población X es más pequeño que en la población Y", gráficamente el escenario es el siguiente:



Entonces:

Bajo H_0 y bajo H_1 tenemos respectivamente :

$$\hat{\theta}_X \approx N(\theta_X, Var(\hat{\theta}_X))$$

$$\hat{\theta}_Y \approx N(\theta_Y, Var(\hat{\theta}_Y))$$

$$P(\hat{\theta}_X > L) = \alpha \quad \text{y} \quad P(\hat{\theta}_Y < L) = \beta$$

por lo tanto :

$$\frac{L-\theta_X}{\sqrt{Var(\hat{\theta}_X)}} = z_{1\text{-}\alpha} \quad \text{y} \quad \frac{L-\theta_Y}{\sqrt{Var(\hat{\theta}_Y)}} = z_\beta$$

de donde :

$$\theta_X + z_{1-\alpha} \cdot \sqrt{Var(\hat{\theta}_X)} = \theta_Y + z_\beta \cdot \sqrt{Var(\hat{\theta}_Y)}$$

$$\theta_X - \theta_Y = z_\beta \cdot \sqrt{Var(\hat{\theta}_Y)} - z_{1-\alpha} \cdot \sqrt{Var(\hat{\theta}_X)}$$

Igualdades de oro :

$$\delta = z_\beta \cdot \sqrt{Var(\hat{\theta}_Y)} - z_{1-\alpha} \cdot \sqrt{Var(\hat{\theta}_X)}$$

si $\theta_X > \theta_Y$

$$\delta = z_\alpha \cdot \sqrt{Var(\hat{\theta}_X)} - z_{1-\beta} \cdot \sqrt{Var(\hat{\theta}_Y)}$$

De estas igualdades de oro despejamos el tamaño de muestra requerido por grupo, que incluso lo supondremos, en forma general, distinto en cada grupo (diseño des balanceado), es decir supondremos que:

$$\frac{n_Y}{n_X} = r \ \ (ratio\ of\ sample\ size), lo\ que\ implica\ \ n_Y = r \cdot n_X$$

Así tememos:

I. El tamaño de muestra para detectar una diferencia de medias:

$$H_0: \mu_X - \mu_y = 0$$
$$H_1: \mu_X - \mu_y = \delta$$

Aquí $\hat{\theta}_X = \bar{X} \rightarrow Var(\hat{\theta}_X) = \frac{\sigma^2}{n_X}$ y $\hat{\theta}_Y = \bar{Y} \rightarrow Var(\hat{\theta}_Y) = \frac{\sigma^2}{n_Y} = \frac{\sigma^2}{r \cdot n_X}$, (suponemos que σ^2 es la misma para la población X y en la población Y) poniéndolos en la primera igualdad de oro, tenemos:

$$\delta = z_\beta \cdot \sqrt{\frac{\sigma^2}{r \cdot n_X}} - z_{1-\alpha} \cdot \sqrt{\frac{\sigma^2}{n_X}} = \frac{\sigma}{\sqrt{n_X}} \cdot \left(z_\beta \cdot \sqrt{\frac{1}{r}} - z_{1-\alpha} \right)$$

De donde:

$$n_X \geq \left(\frac{\sigma}{\delta} \right)^2 \cdot \left(z_\beta \cdot \sqrt{\frac{1}{r}} - z_{1-\alpha} \right)^2, \ n_Y = r \cdot n_X$$

Obviamente que si se diseña un estudio balanceado, r=1, y el tamaño de muestra por grupo es:

$$n_{por\ grupo} \geq \left(\frac{\sigma}{\delta} \right)^2 \cdot \left(z_\beta - z_{1-\alpha} \right)^2$$

Ejemplo: Supongamos que deseamos detectar una diferencia de 5 puntos entre dos grupos (promedio en el grupo 1 es 55 y en el grupo 2 es 50), desviación estándar común es 10, con significación del 5% y potencia del 80%, y:

a) Razón de 2:1 en las muestras:

```
. sampsi 55 50, sd1(10) sd2(10) alpha(0.05) power(.80) ratio(2) onesided

Estimated sample size for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:

         alpha =   0.0500  (one-sided)
         power =   0.8000
            m1 =       55
            m2 =       50
           sd1 =       10
           sd2 =       10
         n2/n1 =     2.00

Estimated required sample sizes:

            n1 =       38
            n2 =       76
```

b) Razón de 1:1 en las muestras:

```
. sampsi 55 50, sd1(10) sd2(10) alpha(0.05) power(.80) ratio(1) onesided

Estimated sample size for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
                    and m2 is the mean in population 2
Assumptions:

         alpha =   0.0500  (one-sided)
         power =   0.8000
            m1 =       55
            m2 =       50
           sd1 =       10
           sd2 =       10
         n2/n1 =     1.00

Estimated required sample sizes:

            n1 =       50
            n2 =       50
```

II. El tamaño de muestra para detectar una diferencia de proporciones:

$$H_0: P_X - P_y = 0$$
$$H_1: P_X - P_y = \delta$$

Con un razonamiento análogo al anterior, para muestras balanceadas se encuentra:

$$n_{por\ grupo} \geq \frac{\left(z_\alpha \sqrt{P_X Q_X} + z_{1-\beta} \sqrt{P_Y Q_Y}\right)^2}{\delta^2}$$

Ejemplo: Se desea probar si el tratamiento A tiene una proporción de mejoría de un 87% versus el placebo que tiene una proporción de mejoría del 65%. ¿Cuántos pacientes habrá que aleatorizar por grupo? Si se desea una significación del 1% y una potencia del 99%

```
. sampsi .87 .65, alpha(0.01) power(.99) onesided

Estimated sample size for two-sample comparison of proportions

Test Ho: p1 = p2, where p1 is the proportion in population 1
                    and p2 is the proportion in population 2
Assumptions:

         alpha =   0.0100  (one-sided)
         power =   0.9900
            p1 =   0.8700
            p2 =   0.6500
         n2/n1 =   1.00

Estimated required sample sizes:

            n1 =      167
            n2 =      167
```

Dócimas de normalidad para variables contínuas

Ya hemos dicho que el test t-Student, se sustenta sobre la normalidad de la variable que se desea docimar, por lo tanto es de sumo interés saber si este supuesto puede ser descartado o no, con este propósito, se introducen las dócimas de normalidad. En todas ellas la hipótesis nula es que la variable en estudio tiene una distribución normal. Stata provee de algunos test para este efecto, los más conocidos son:

- El test de Shapiro-Wilk
- El test de Shapiro-Francia
- El S-K test

Ejemplo: Usando la base de datos stress_cortisol.dta, docimar la normalidad de la variable edad:

```
. swilk  edad

                   Shapiro-Wilk W test for normal data
    Variable |    Obs        W          V          z     Prob>z
-------------+-------------------------------------------------
        edad |    500    0.99883      0.395     -2.234  0.98725

. sfrancia  edad

                  Shapiro-Francia W' test for normal data
    Variable |    Obs        W'         V'         z     Prob>z
-------------+-------------------------------------------------
        edad |    500    0.99886      0.409     -2.053  0.97998

. sktest  edad

                   Skewness/Kurtosis tests for Normality
                                                 ------- joint ------
    Variable |  Pr(Skewness)   Pr(Kurtosis)  adj chi2(2)    Prob>chi2
-------------+-------------------------------------------------------
        edad |      0.681         0.565           0.50       0.7774
```

Nociones de inferencia no paramétrica

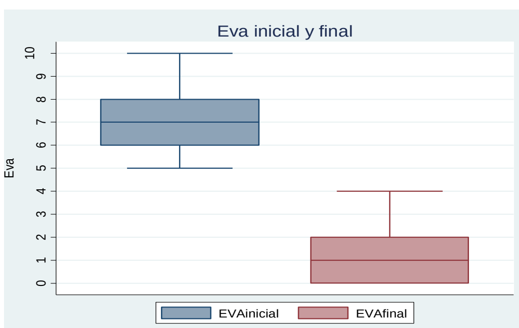
¿Qué ocurre cuando el supuesto de normalidad no se satisface y la muestra es pequeña? O cuándo la variable no es de naturaleza numérica y es sólo ordinal. En estos casos está la alternativa no paramétrica, que la presentaremos en una secuencia de ejemplos.

1. Consideremos la siguiente situación: Se desea analizar un estudio clínico para comparar la respuesta de un opiode ante dolor de naturaleza oncológica. El diseño es antes-después, es decir se "mide" la sensación de dolor antes de administrar el tratamiento y luego de ser administrado. La cuantificación de la sensación de dolor fue medida a través de la escala EVA. Además se necesita comparar la respuesta según sexo del paciente.

```
Los datos se encuentran en la base EVA.dta
  obs:            30
 vars:             4                          4 Jan 2006 09:53
 size:           510 (99.9% of memory free)
-------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
pac             float  %9.0g
EVAinicial      float  %9.0g
EVAfinal        float  %9.0g
sexo            byte   %8.0g                  0:hombres 1: mujeres
-------------------------------------------------------------------------------
Sorted by:  pac
```

Estadísticas descriptivas:

```
. tabstat  EVAinicial EVAfinal, stat(n min q max) col(stat)
    variable |         N       min       p25       p50       p75       max
-------------+------------------------------------------------------------
  EVAinicial |        30         5         6         7         8        10
    EVAfinal |        30         0         0         1         2         4
--------------------------------------------------------------------------

graph box EVAinicial EVAfinal, ytitle(Eva) ylabel(0(1)10) title(Eva inicial y final)
```



Evaluación del cambio de EVA:

```
. gen difEVA= EVAfinal- EVAinicial

. tabstat   difEVA, stat(n min q max) col(stat)
    variable |         N       min       p25       p50       p75       max
-------------+------------------------------------------------------------
      difEVA |        30       -10        -7        -6        -4        -1
--------------------------------------------------------------------------
```

Para evaluar el cambio de EVA se usa el test de Wilcoxon

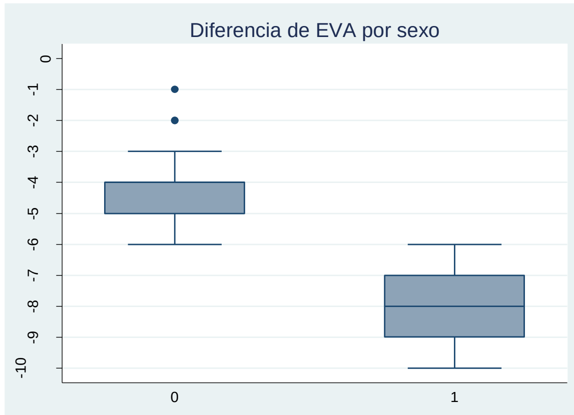
```
. signrank  EVAinicial= EVAfinal
Wilcoxon signed-rank test
        sign |      obs   sum ranks    expected
-------------+---------------------------------
    positive |       30         465       232.5
    negative |        0           0       232.5
        zero |        0           0           0
-------------+---------------------------------
         all |       30         465         465
unadjusted variance     2363.75
adjustment for ties      -10.88
adjustment for zeros       0.00
                     ----------
adjusted variance       2352.88
Ho: EVAinicial = EVAfinal
             z =   4.793
    Prob > |z| =   0.0000
```

```
. signrank   difEVA=0
Wilcoxon signed-rank test
        sign |      obs   sum ranks    expected
-------------+---------------------------------
    positive |        0           0       232.5
    negative |       30         465       232.5
        zero |        0           0           0
-------------+---------------------------------
         all |       30         465         465
unadjusted variance     2363.75
adjustment for ties      -10.88
adjustment for zeros       0.00
                     ----------
adjusted variance       2352.88
Ho: difEVA = 0
             z =  -4.793
    Prob > |z| =   0.0000
```

El cambio de EVA por sexo

```
. signrank   difEVA=0
Wilcoxon signed-rank test
        sign |      obs   sum ranks    expected
-------------+---------------------------------
    positive |        0           0       232.5
    negative |       30         465       232.5
        zero |        0           0           0
-------------+---------------------------------
         all |       30         465         465
unadjusted variance     2363.75
adjustment for ties      -10.88
adjustment for zeros       0.00
                     ----------
adjusted variance       2352.88
Ho: difEVA = 0
             z =  -4.793
    Prob > |z| =   0.0000
```

```
graph box difEVA, over(sexo) title(Diferencia de EVA por sexo ) ytitle(Diferencia de EVA) ylabel(-10(1)0)
```



Para evaluar el cambio de EVA por sexo se usa el test de Mann-Whitney

```
. ranksum  difEVA, by( sexo)
Two-sample Wilcoxon rank-sum (Mann-Whitney) test
        sexo |      obs    rank sum    expected
-------------+---------------------------------
           0 |       17         370       263.5
           1 |       13          95       201.5
-------------+---------------------------------
    combined |       30         465         465
unadjusted variance      570.92
adjustment for ties      -11.05
                     ----------
adjusted variance        559.87
Ho: difEVA(sexo==0) = difEVA(sexo==1)
             z =   4.501
    Prob > |z| =   0.0000
```

- Dócima de independencia entre dos variables nominales:

Con frecuencia queremos averiguar si dos variables cualitativas X e Y están vinculadas. En cada unidad de observación se registra un par (x,y) de valores observados, en consecuencia a partir de lo obtenido en n unidades de observación, se obtiene una Tabla de Contingencia de s×r (tabla observada):

	X_1	X_2	...X_r	*Total*
Y_1	O_{11}	O_{12}	O_{1r}	$n_{1.}$
Y_2	O_{21}	O_{22}	O_{2r}	$n_{2.}$
...Y_s	O_{s1}	O_{s2}	O_{sr}	$n_{s.}$
Total	$n_{.1}$	$n_{.2}$	$n_{.r}$	n

Bajo la Hipótesis de independencia, estas frecuencias se pueden recalcular, creándose una Tabla Esperada:

	X_1	X_2	...X_r	*Total*
Y_1	E_{11}	E_{12}	E_{1r}	
Y_2	E_{21}	E_{22}	E_{2r}	
...Y_s	E_{s1}	E_{s2}	E_{sr}	
Total				n

Donde :

$$E_{ij} = \frac{n_{.j} \cdot n_{i.}}{n}$$

En estas condiciones, podemos plantear la Hipótesis Nula:

H_0 : X es independiente de Y

Contrastada con la Hipótesis alternativa:

H_1 : X está asociada con Y

La estadística de prueba es:

$$\sum_{ij} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

que sigue una distribución $\chi^2_{((s-1)\cdot(r-1))}$

Si a lo menos una de las celdas de la tabla esperada, contiene una frecuencia menor a 5, es posible que la estadística de Chi cuadrado se haga muy grande, sólo por el hecho que la frecuencia esperada es pequeña y por lo tanto asociar un p-value pequeño y sin embargo no haber asociación. Para solucionar este problema, tenemos como alternativa a la dócima de chi cuadrado el Test Exacto de Fisher:

```
. tabi 40 3 \ 1 1, chi2 exact expected

+--------------------+
| Key                |
|--------------------|
|     frequency      |
| expected frequency |
+--------------------+

           |          col
       row |         1          2 |     Total
-----------+----------------------+----------
         1 |        40          3 |        43
           |      39.2        3.8 |      43.0
-----------+----------------------+----------
         2 |         1          1 |         2
           |       1.8        0.2 |       2.0
-----------+----------------------+----------
     Total |        41          4 |        45
           |      41.0        4.0 |      45.0

          Pearson chi2(1) =   4.3679   Pr = 0.037 *
           Fisher's exact =                 0.172 **
   1-sided Fisher's exact =                 0.172
```

- Según Chi cuadrado las filas y columnas están asociadas.
- Según el test exacto de Fisher no lo están

Dado que hay frecuencias esperadas muy pequeñas, el test exacto de Fisher es mas confiable.

Un caso particular de esta dócima es la clásica tabla tetracórica:

	Enfermo	Sano
Expuesto	A	B
No expuesto	C	D

La única novedad, que en ella, dependiendo del diseño utilizado al conseguirla, podemos definir las medidas de asociación vistas en la primera parte del capítulo de probabilidades:

- El riesgo relativo:

$$RR = \frac{\frac{A}{A+B}}{\frac{C}{C+D}} = \frac{A(C+D)}{C(A+B)}$$

```
. csi 40 3  1 1

                 |   Exposed   Unexposed  |      Total
-----------------+------------------------+------------
           Cases |        40           3  |         43
        Noncases |         1           1  |          2
-----------------+------------------------+------------
           Total |        41           4  |         45
                 |                        |
            Risk |  .9756098         .75  |   .9555556
                 |                        |
                 |      Point estimate    |    [95% Conf. Interval]
                 |------------------------+------------------------
 Risk difference |         .2256098       |   -.2013538    .6525733
      Risk ratio |         1.300813       |    .7372195    2.295266
 Attr. frac. ex. |           .23125       |   -.3564481    .5643206
 Attr. frac. pop |         .2151163       |
                 +-------------------------------------------------
                               chi2(1) =     4.37  Pr>chi2 = 0.0366
```

- El Odds Ratio:

$$OR = \frac{\frac{A}{B}}{\frac{C}{D}} = \frac{AD}{BC}$$

```
. cci 40 3  1 1
                                                         Proportion
                 |   Exposed   Unexposed  |      Total     Exposed
-----------------+------------------------+------------------------
           Cases |        40           3  |         43       0.9302
        Controls |         1           1  |          2       0.5000
-----------------+------------------------+------------------------
           Total |        41           4  |         45       0.9111
                 |                        |
                 |      Point estimate    |    [95% Conf. Interval]
                 |------------------------+------------------------
      Odds ratio |         13.33333       |    .1276094    1070.974 (exact)
 Attr. frac. ex. |             .925       |   -6.836416    .9990663 (exact)
 Attr. frac. pop |         .8604651       |
                 +-------------------------------------------------
                               chi2(1) =     4.37  Pr>chi2 = 0.0366
```

El test de Mantel-Haenszel

Supongamos que en un estudio transversal basado en una muestra aleatoria poblacional, se pregunta por la condición de diabético tipo 2, presencia de síndrome metabólico y sexo, se hace el siguiente análisis:

```
. cc   SM DIABETES
                                                         Proportion
                 |   Exposed   Unexposed  |      Total     Exposed
-----------------+------------------------+------------------------
           Cases |       117         341  |        458       0.2555
        Controls |        75        1113  |       1188       0.0631
-----------------+------------------------+------------------------
           Total |       192        1454  |       1646       0.1166
                 |                        |
                 |      Point estimate    |    [95% Conf. Interval]
                 |------------------------+------------------------
      Odds ratio |         5.09173        |    3.675926    7.068656 (exact)
 Attr. frac. ex. |        .8036031        |    .7279597    .8585304 (exact)
 Attr. frac. pop |        .2052873        |
                 +-------------------------------------------------
                               chi2(1) =   118.67  Pr>chi2 = 0.0000
```

Al estratificar por sexo:

```
. cc   SM DIABETES if sexo==0
                                                         Proportion
                 |   Exposed   Unexposed  |      Total     Exposed
-----------------+------------------------+------------------------
           Cases |        48         119  |        167       0.2874
        Controls |        28         485  |        513       0.0546
-----------------+------------------------+------------------------
           Total |        76         604  |        680       0.1118
                 |                        |
                 |      Point estimate    |    [95% Conf. Interval]
                 |------------------------+------------------------
      Odds ratio |        6.986795        |     4.08771    12.03549 (exact)
 Attr. frac. ex. |        .8568729        |    .7553643    .9169124 (exact)
 Attr. frac. pop |        .2462868        |
                 +-------------------------------------------------
                               chi2(1) =    68.81  Pr>chi2 = 0.0000
```

```
. cc   SM DIABETES if sexo==1
                                                         Proportion
                 |   Exposed   Unexposed  |      Total     Exposed
-----------------+------------------------+------------------------
           Cases |        69         222  |        291       0.2371
        Controls |        47         628  |        675       0.0696
-----------------+------------------------+------------------------
           Total |       116         850  |        966       0.1201
                 |                        |
                 |      Point estimate    |    [95% Conf. Interval]
                 |------------------------+------------------------
      Odds ratio |        4.152961        |     2.72903    6.341228 (exact)
 Attr. frac. ex. |         .759208        |    .6335695    .8423018 (exact)
 Attr. frac. pop |        .1800184        |
                 +-------------------------------------------------
                               chi2(1) =    53.98  Pr>chi2 = 0.0000
```

A la luz de estos resultados, ¿hay evidencia para creer en la interacción con sexo? Este problema lo resuelve el test de Mantel-Haenszel, cuya hipótesis nula es la “homogeneidad” de los OR:

```
. cc   SM DIABETES, by(sexo)

0: Hombre 1: Muj |       OR       [95% Conf. Interval]   M-H Weight
-----------------+-------------------------------------------------
               0 |   6.986795       4.08771   12.03549          4.9 (exact)
               1 |   4.152961       2.72903   6.341228     10.80124 (exact)
-----------------+-------------------------------------------------
           Crude |    5.09173      3.675926   7.068656              (exact)
     M-H combined |   5.037336      3.679703   6.895871
-------------------------------------------------------------------
Test of homogeneity (M-H)      chi2(1) =     2.49  Pr>chi2 = 0.1148

                   Test that combined OR = 1:
                                Mantel-Haenszel chi2(1) =    118.07
                                                Pr>chi2 =    0.0000
```

El p-value en negritas indica que no hay suficiente evidencia para creer que el riesgo de presencia de diabetes para el síndrome metabólico sea distinto por sexo.