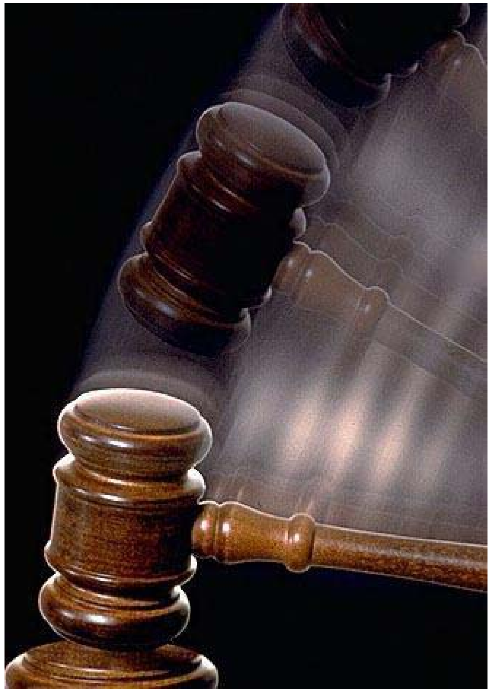
Tema 11 Test de hipótesis (I)

1. **Conceptos básicos**
2. **Hipótesis simple. Lema de Neyman-Pearson.**
3. **Hipótesis compuesta.**
    1. **Cociente de verosimilitud.**
    2. **Distribución asintótica de λ**
    3. **Familias paramétricas diferentes**
    4. **Ejemplo. Búsquedas**
4. Test de paramétricos para variables normales
5. Test de bondad de ajuste *(Goodness-of-fit)*
    1. Test de χ^2 de Pearson
    2. Test de Kolmogorov-Smirnov
6. Test de consistencia y aleatoriedad
    1. Test Run
    2. Test de Kolmogorov-Smirnov para dos muestras
7. Test de independencia



1. Conceptos básicos

Formulamos una hipótesis de forma precisa sobre un experimento → Adquirimos datos →
La idea es utilizar los datos experimentales para aceptar o rechazar la hipótesis
¿¡En caso de error!?

Objetivo.- **Reducir las posibilidades de error (equivocación) al mínimo aceptable**

Una hipótesis es una afirmación que puede ser probada experimentalmente

Tipos de hipótesis

Simple .- Todos los parámetros de la pdf de la variable aleatoria (datos) están especificados
e.g. los valores se distribuyen normalmente con media μ *y anchura* σ

Compuesta .- No se proporcionan todos los parámetros de la pdf
e.g. los valores se distribuyen normalmente con media $\mu > 5$

Si la pdf tiene n parámetros,
una hipótesis simple selecciona un único punto en un espacio n-dimensional,
una hipótesis compuesta selecciona un sub-espacio.

1. Conceptos básicos

Hipótesis nula H_0 (*Null Hipothesis)*

Se define como la hipótesis que estamos considerando, que queremos testear, H_0

En general, cualquier afirmación sobre H_0 se realiza con referencia a hipótesis alternativas, H_1, H_2,…

Supongamos que tenemos n medidas independientes $\overline{x} = \{x_1, x_2, \ldots, x_n\}$ de la variable aleatoria x y un conjunto de hipótesis: H_0, H_1, H_2,… cada una de ellas especifica una pdf dada:

$$f\left(\overline{x}\middle|H_0\right), f\left(\overline{x}\middle|H_1\right), f\left(\overline{x}\middle|H_2\right)$$

En estadística uno no puede **aceptar** una hipótesis: estrictamente **SOLO no rechazarla** o **rechazarla**

No decimos que hay una señal, **SINO** que podemos excluir la presencia de una señal.

En general: $H_0 \rightarrow$ Solo *background* $H_1 \rightarrow$ Presencia de señal

Test estadístico

Para investigar el acuerdo entre los datos observados y una hipótesis dada se construye una función de las variables medidas $\overline{x} = \{x_1, x_2, \ldots, x_n\}$ y de la hipótesis:

TEST ESTADÍSTICO $t\left(\overline{x}\right)$

Puede ser unidimensional o multidimensional

Cada una de las hipótesis : H_0, H_1, H_2,… especifica una pdf "*muy diferente*" para el estadístico $t\left(\overline{x}\right)$

$$g\left(t\middle|H_0\right), g\left(t\middle|H_1\right), g\left(t\middle|H_2\right)$$

1. Conceptos básicos

Tipos de errores

Tipo I : Rechazar una hipótesis verdadera (Falso negativo)

Tipo II : Aceptar una hipótesis falsa (Falso positivo)

Ejemplos:

Tribunal de justicia: "El acusado se proclama inocente" (H_0)

Tipo I : Realmente es inocente pero el jurado le condena.

Tipo II : Realmente es culpable (miente bien) y el tribunal lo declara inocente.

Búsqueda de partículas: (H_0 es presencia de fondo solamente)

Tipo I : Realmente no hay señal, pero publicamos la presencia de una resonancia.

Tipo II : Sí había resonancia, pero aceptamos H_0 y perdemos el Nobel.

1. Conceptos básicos

Región crítica y región de aceptancia

La compatibilidad entre **Hipótesis** y **Datos** se plantea en términos de aceptar o rechazar la hipótesis nula H_0

El espacio de valores del test estadístico, $t(\overline{x})$, se divide en dos regiones o sub-espacios:

Región crítica : Si el valor de t cae en esta región → Rechazamos H_0

Región de aceptancia: Si cae en esta región → Aceptamos H_0

Significancia (*Size)*

Probabilidad de cometer un error de **Tipo I**

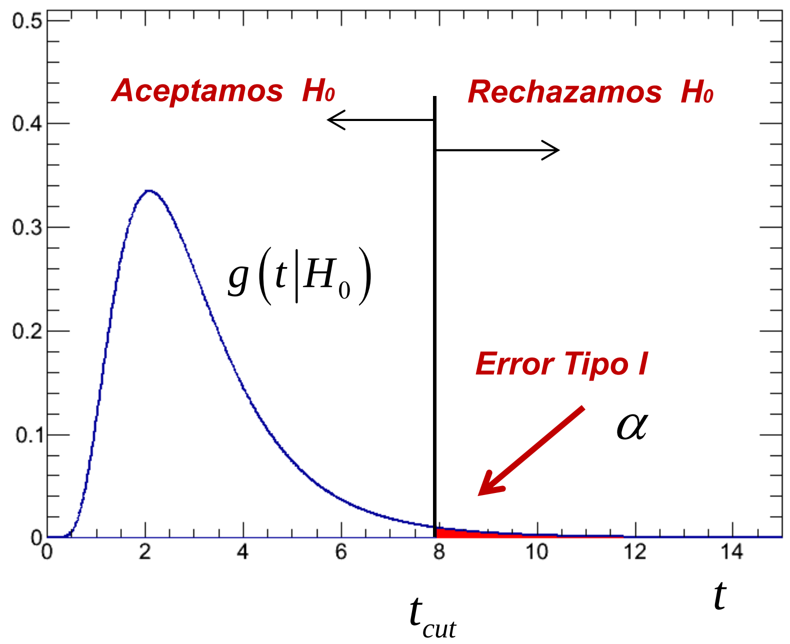
$$\alpha = \int_{t_{cut}}^{+\infty} g\left(t \middle| H_0\right) dt$$

Hay una probabilidad α de rechazar H_0 si H_0 es cierta.

Valores típicos: 5%, 1%

Un test es significativo al nivel $1-\alpha$ si la probabilidad de rechazar la hipótesis verdadera es menor o igual que α

Hipótesis H0



1. Conceptos básicos

Potencia del test (Power)

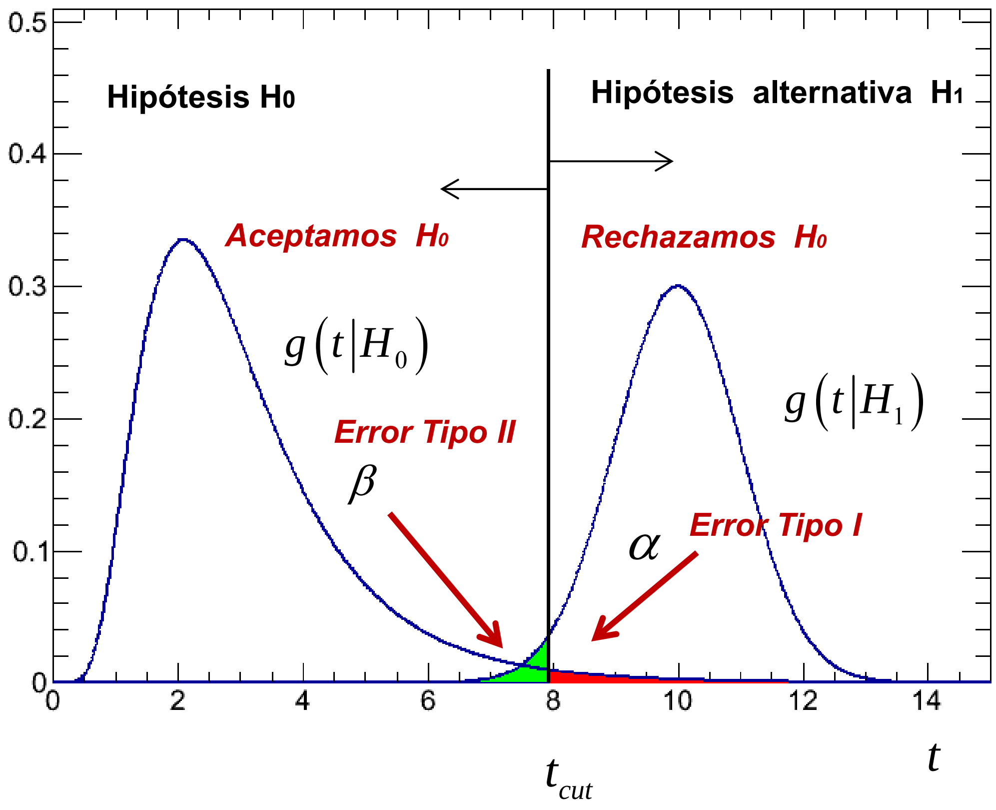
Supongamos que tenemos una hipótesis alternativa H_1 y conocemos: $g(t|H_1)$

La probabilidad de cometer un error de **Tipo II** , es decir de aceptar la hipótesis nula H_0 cuando H_1 es cierta viene dada por:

$$\beta = \int_{-\infty}^{t_{cut}} g(t|H_1)\,dt$$

A la cantidad $1-\beta$ se le denomina **potencia** del test o capacidad de discriminar la hipótesis alternativa H_1

Claramente, cuanto menor sean los valores de α y β mejor y más potente es el test (i.e. H_0 y H_1 muy diferentes entre si).



1. Conceptos básicos

Frecuentemente, lo que queremos es medir el nivel de compatibilidad entre los datos y la hipótesis H_0 sin hacer referencia a hipótesis alternativas

¿Hasta que punto H_0 es compatible con los datos?

¿La función que estamos usando describe bien los datos?

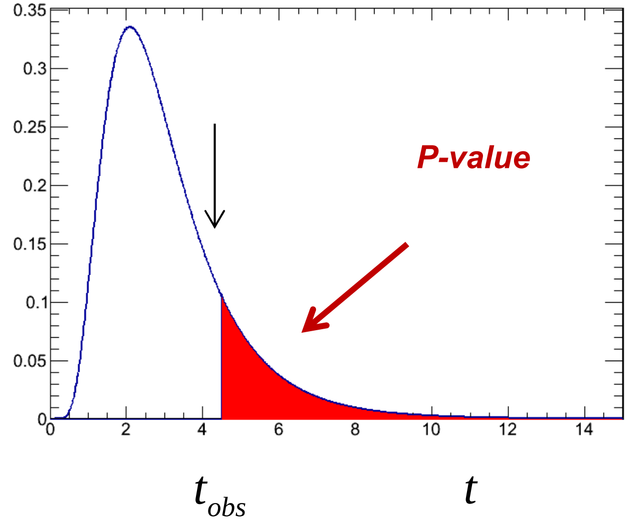
P-value

El nivel de compatibilidad entre datos e hipótesis se puede cuantificar en términos del p-value

P-value .- Probabilidad de que, asumiendo como cierta la hipótesis en cuestión H_0, observemos un valor del test estadístico igual o menos compatible con H_0 que el obtenido.

¡Atención! el ***p-value*** **NO** es la probabilidad de que "**H_0 sea cierta**".

Los ***p-values*** se distribuyen uniformemente si la hipótesis es cierta



¿Tiene sentido preguntar cual es la probabilidad de observar un valor dado $t = t_{obs}$?

La región crítica viene definida por un $p_{crit} = \alpha$

Si $p_{crit} = 0.05$ y $p < p_{crit}$ ➡ Hipótesis excluida al 95 % CL.

1. Conceptos básicos

Ejemplo: Test de hipótesis frecuentista

PASO 1 Diseñar una cantidad (test estadístico) que dependa de los datos y los clasifique según sean señal o fondo.

Ejemplo: buscar una nueva partícula contando sucesos que pasan un criterio de selección

H_0 : "*Solo hay fondo*" esperamos *b* sucesos

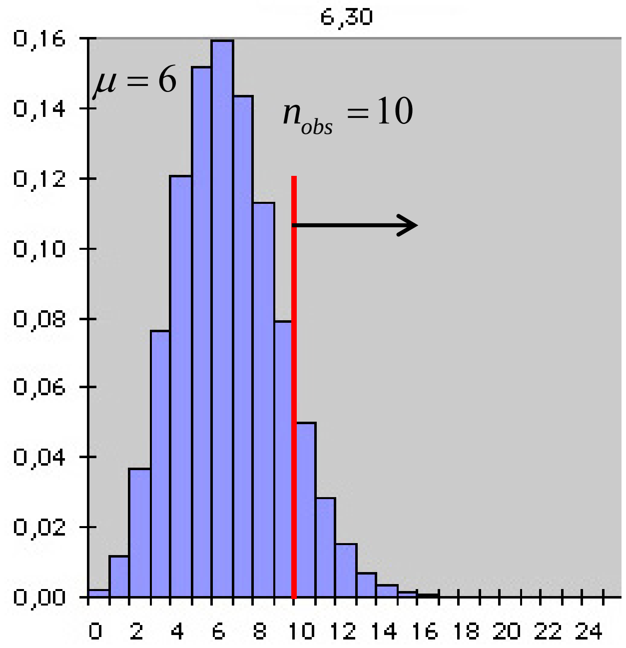
H_1 : "*Fondo más señal*" esperamos *s+b* sucesos

Test estadístico n_{obs}

PASO 2 Calcular la distribución del test estadístico si H_0 y si H_1. Dos distribuciones

PASO 3 Realizar el experimento y obtener el valor del test estadístico: n_{obs}

PASO 4 Calcular el p-value: $p(n \geq n_{\text{obs}} | H_0)$



Ejemplo: H_0 : "*Solo hay fondo*" $b = \mu = 6$ $n_{obs} = 10$

p-value = 0.084

$$P(n|b) = \frac{b^n}{n!} e^{-b} \qquad P(n|s+b) = \frac{(s+b)^n}{n!} e^{-(s+b)}$$

Si $p_{crit} = 0.05$ $p > p_{crit}$ ➡ Hipótesis aceptada

1. Conceptos básicos

Los físicos hablan de los resultados en términos de “**sigmas**” en lugar de “**p-values**”

El número de sigmas es el “z-value” correspondiente a una distribución normal estándar

Por convenio:

Rechazar H_0 con $z > 3\,\sigma$ (i.e. a 3σ) se denomina una “***EVIDENCE***”

Rechazar H_0 con $z > 5\,\sigma$ (i.e. a 5σ) se denomina una “***DISCOVERY***”

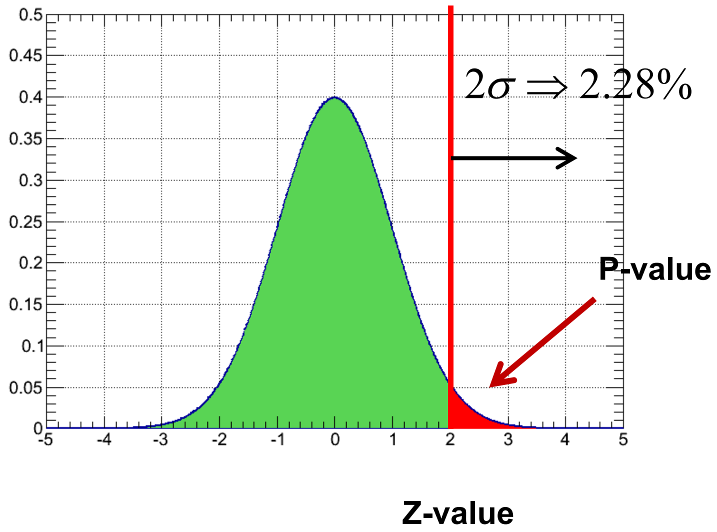
ROOT Functions

```
Double_t  zvalue = - TMath::NormQuantile (Double_t  pvalue)

Double_t  pvalue =  1 -  TMath::Freq (Double_t z-value)
```

$$\text{p-value} = \frac{\left(1 - \mathrm{Erf}\left(\frac{\text{zvalue}}{\sqrt{2}}\right)\right)}{2}$$

Z-value (σ)	P-value
1.0	0.159
2.0	0.0228
3.0	0.00135
4.0	3.17×10^{-5}
5.0	2.87×10^{-7}



2. Hipótesis simples. Lema de Neyman-Pearson

El mejor test es aquel para el que tanto α como β son lo más pequeño posible.
Solo se puede encontrar si y solo si ambas hipótesis H_0 y H_1 son hipótesis simples

Supongamos una variable aleatoria *x* tal que:

- Bajo H_0 se distribuye como: $f\left(x|H_0\right) = f\left(x|\theta_0\right)$
- Bajo H_1 se distribuye como: $f\left(x|H_1\right) = f\left(x|\theta_1\right)$

Lema de Neyman-Pearson

"Para un nivel de significancia α dado, la región crítica w_α para la que se obtiene la mayor potencia (1- β) respecto a la hipótesis alternativa H_1, es aquella que incluye los valores de x para los cuales $f\left(x|H_1\right)$ es tan grande como sea posible respecto a $f\left(x|H_0\right)$"

Demostración

Tamaño o nivel de significancia del test

$$P\left(x \in w_\alpha \middle| H_0\right) = \int_{w_\alpha} f\left(x|\theta_0\right) dx = \alpha$$

Potencia del test respecto a H_1

$$P\left(x \in w_\alpha \middle| H_1\right) = \int_{w_\alpha} f\left(x|\theta_1\right) dx = 1 - \beta$$

Buscamos la región crítica w_α que maximiza potencia (1- β) para un α dado

2. Hipótesis simples. Lema de Neyman-Pearson

Demostración

Teorema del valor medio ponderado para integrales

$$1-\beta = \int_{w_\alpha} f\left(x|\theta_1\right)dx = \int_{w_\alpha} \frac{f\left(x|\theta_1\right)}{f\left(x|\theta_0\right)} f\left(x|\theta_0\right)dx = \left.\frac{f\left(x|\theta_1\right)}{f\left(x|\theta_0\right)}\right|_{x=\xi} \int_{w_\alpha} f\left(x|\theta_0\right)dx = \left.\frac{f\left(x|\theta_1\right)}{f\left(x|\theta_0\right)}\right|_{x=\xi} \alpha$$

$$\frac{1-\beta}{\alpha} = \left.\frac{f\left(x|\theta_1\right)}{f\left(x|\theta_0\right)}\right|_{x=\xi}$$

La mejor región crítica es la que hace que el cociente sea máximo

$$\frac{f\left(x|\theta_1\right)}{f\left(x|\theta_0\right)} > k$$

Donde k viene determinado por el nivel de confianza α

Para un conjunto de medidas o datos:

$$L\left(\overline{x}|\theta\right) = \prod_{i=1}^{n} f\left(x_i|\theta_0\right)$$

$$\frac{L\left(\overline{x}|\theta_1\right)}{L\left(\overline{x}|\theta_0\right)} = \frac{L\left(\overline{x}|H_1\right)}{L\left(\overline{x}|H_0\right)} > k$$

- La condición anterior proporciona la región crítica con la máxima potencia de la hipótesis H_0 frente a H_1
- Solo se puede aplicar cuando ambas hipótesis son **simples**
- En la práctica, la región puede definirse a partir del estadístico unidimensional: $t = \dfrac{L\left(\overline{x}|H_1\right)}{L\left(\overline{x}|H_0\right)}$

2. Hipótesis simples. Lema de Neyman-Pearson

Ejemplo: Vida media

$$f\left(t \mid \tau\right) = \frac{1}{\tau} e^{-t/\tau}$$

$H_0 : \tau = 1$

$H_1 : \tau = 2$

$$\frac{L\left(t \mid H_1\right)}{L\left(t \mid H_0\right)} = \frac{\prod \frac{1}{2} e^{-t_i/2}}{\prod e^{-t_i/1}} = \left(\frac{1}{2}\right)^n \exp\left[\frac{1}{2}\sum t_i\right] > k$$

En términos de la media:

$$\frac{L\left(t \mid H_1\right)}{L\left(t \mid H_0\right)} = \left(\frac{1}{2}\right)^n \exp\left[\frac{n\bar{t}}{2}\right] > k \quad \Rightarrow \quad \frac{1}{2} n\bar{t} > \ln\left(k 2^n\right)$$

$$\boxed{\bar{t} > 2\left(\frac{\ln k}{n} + \ln 2\right) = T_n}$$

La mejor región crítica depende de T_n que a su vez depende de α

Para conocer la región crítica necesitamos conocer la pdf del estadístico $f\left(\bar{t} \mid H_0\right)$. Supongamos α = 5 %

CASO A: *n = 1*

$n = 1$

$\bar{t} = t$

$f(t) = \frac{1}{\tau} e^{-t/\tau}$

$$0.05 = \alpha = \int_{w_\alpha} f\left(\bar{t} \mid H_0\right) d\bar{t} = \int_{T_n}^{\infty} e^{-\bar{t}} d\bar{t} = e^{-T_1} \quad \Rightarrow \quad T_1 = -\ln\alpha = 3.00$$

Si $\bar{t}_{obs} > T_1 = 3.00$ se rechaza H_0

Potencia del test

$$1 - \beta = \int_{w_\alpha} f\left(\bar{t} \mid H_1\right) d\bar{t} = \int_{T_n}^{\infty} \frac{1}{2} e^{-\frac{\bar{t}}{2}} d\bar{t} = e^{-\frac{T_1}{2}} = 0.22$$

Si H_1 es cierto el 78 % de las veces cometemos un error tipo II

2. Hipótesis simples. Lema de Neyman-Pearson

Ejemplo: Vida media

$$f\left(t\middle|\tau\right)=\frac{1}{\tau}e^{-t/\tau}$$

CASO B: *n* muy grande →

$$f\left(\bar{t}\right)=N\left(\tau,\frac{\tau^2}{n}\right)=\frac{1}{\sqrt{2\pi}\,\tau/\sqrt{n}}\exp\left(-\frac{1}{2}\frac{\left(\bar{t}-\tau\right)^2}{\tau^2/n}\right)$$

→

$$0.05=\alpha=\int_{w_\alpha} f\left(\bar{t}\middle|H_0\right)d\bar{t}=\int_{T_n}^{\infty} N\left(1,\frac{1}{n}\right)d\bar{t}=1-\Phi\left(\frac{T_n-1}{1/\sqrt{n}}\right)$$

→

$$T_n=1+\frac{\Phi^{-1}\left(0.95\right)}{\sqrt{n}}=1+\frac{1.645}{\sqrt{n}}$$

→ Depende de n:

$$n=100 \qquad T_{100}=1.1645$$

→ Si $\bar{t}_{obs}>T_{100}=1.1645$ se rechaza H_0

Potencia del test

$$1-\beta=\int_{w_\alpha} f\left(\bar{t}\middle|H_1\right)d\bar{t}=\int_{T_n}^{\infty} N\left(2,\frac{4}{n}\right)d\bar{t}=1-\Phi\left(\frac{T_n-2}{2/\sqrt{n}}\right)$$

Depende de n también:

Para $n=100$

$1-\beta=0.99998$

La potencia de H_0 respecto a H_1 aumenta muy rápido con el número de observaciones

3. Hipótesis compuestas. Test cociente de verosimilitud

El lema de Neyman-Pearson solo se puede aplicar si ambas hipótesis H_0 y H_1 son hipótesis simples
Cuando se trata de hipótesis compuestas se suele aplicar el test de ***cociente de verosimilitudes***:
(likelihood ratio test)

Distinguimos dos situaciones posibles:

- **Familias paramétricas idénticas**.- Tanto para H_0 como H_1 las pdf's pertenecen a una familia continua de hipótesis.

  Ejemplo:
  $$H_0: \quad f(x:\theta) \quad \text{con} \quad \theta < \theta_0$$
  $$H_1: \quad f(x:\theta) \quad \text{con} \quad \theta > \theta_0$$

- **Familias paramétricas diferentes**.- Las pdf's de H_0 como H_1 pertenecen a familias diferentes de hipótesis.
  $$H_0: \quad f(x:\overline{\phi}) \quad \text{con} \quad \phi \text{ parámetros libres (sin especificar)}$$
  $$H_1: \quad g(x:\overline{\psi}) \quad \text{con} \quad \psi \text{ parámetros libres (sin especificar)}$$

3.1 Test cociente de verosimilitud

Propuesto por Neyman en 1928

Supongamos n observaciones $\overline{x} = \{x_1, x_2, \ldots, x_n\}$ independientes y que ambas hipótesis especifican la misma pdf $f\left(\overline{x}; \overline{\theta}\right)$

Las hipótesis pueden especificarse como:

Ω .- Espacio de parámetros

w .- Subespacio $w \subset \Omega$

$$H_0: \quad \overline{\theta} \in w$$
$$H_1: \quad \overline{\theta} \in \Omega - w$$

Función de verosimilitud $L\left(\overline{x}; \overline{\theta}\right) = \prod_{i=1}^{n} f\left(x_i; \overline{\theta}\right)$

Ejemplos:

$H_0: \quad \theta_1 = a; \quad \theta_2 = b;$	$H_0: \quad \theta_1 = c; \quad \theta_2$ sin especificar	$H_0: \quad \theta_1 + \theta_2 = d;$
$H_1: \quad \theta_1 \neq a; \quad \theta_2 \neq b;$	$H_1: \quad \theta_1 \neq c; \quad \theta_2$ sin especificar	$H_1: \quad \theta_1 + \theta_2 \neq d;$

$L_{w,\max}\left(\overline{x}; \overline{\theta}\right)$ **Máxima verosimilitud condicional**
Máxima verosimilitud en el sub-espacio w

$L_{\Omega,\max}\left(\overline{x}; \overline{\theta}\right)$ **Máxima verosimilitud incondicional**
Máxima verosimilitud en todo el espacio

Se define el cociente de verosimilitudes como:

$$\lambda = \frac{L_{w,\max}\left(\overline{x}; \overline{\theta}\right)}{L_{\Omega,\max}\left(\overline{x}; \overline{\theta}\right)}$$

3.1 Test cociente de verosimilitud

$$\lambda = \frac{L_{w,\max}\left(\overline{x};\overline{\theta}\right)}{L_{\Omega,\max}\left(\overline{x};\overline{\theta}\right)}$$

Es no negativo y cumple: $0 \le \lambda \le 1$

Es función de los observables

$\lambda \to 1$: H0 es más probable

$\lambda \to 0$: H0 es más improbable

Región crítica: $0 \le \lambda \le \lambda_\alpha$

Importante: Necesitamos conocer la pdf para λ suponiendo que H0 sea cierta:

$$g\left(\lambda | H_0\right) \quad \Rightarrow \quad \alpha = \int_0^{\lambda_\alpha} g\left(\lambda | H_0\right) d\lambda$$

Se trata de un test con buenas propiedades, especialmente en el límite de grandes números

Notación más compacta:

$$\left.\begin{array}{llll} H_0: & \theta_i = \theta_{i0}; & i = 1,2,\cdots,r & \overline{\theta}_r = \overline{\theta}_{r0} \\ & \theta_j \quad \text{sin especificar;} & j = 1,2,\cdots,s & \overline{\theta}_s \\ & & & \\ H_1: & \theta_i \ne \theta_{i0}; & i = 1,2,\cdots,r & \overline{\theta}_r \ne \overline{\theta}_{r0} \\ & \theta_j \quad \text{sin especificar;} & j = 1,2,\cdots,s & \overline{\theta}_s \end{array}\right\}$$

$\widehat{\widehat{\overline{\theta}}}_s \to$ Mejor estimación en el subespacio w

$\widehat{\overline{\theta}}_r, \widehat{\overline{\theta}}_s \to$ Mejor estimación en el espacio total Ω

$$\lambda = \frac{L\left(\overline{x};\overline{\theta}_{r0},\widehat{\widehat{\overline{\theta}}}_s\right)}{L\left(\overline{x};\widehat{\overline{\theta}}_r,\widehat{\overline{\theta}}_s\right)}$$

$\overline{\theta}_r \to$ Parámetros especificados

$\overline{\theta}_s \to$ Parámetros sin especificar ***(“nuisance parameters”)***

3.1 Test cociente de verosimilitud

$\hat{\hat{\theta}}_s$ Doble gorro significa que, dados los parámetros especificados $\bar{\theta}_{r0}$ maximizamos la verosimilitud solo respecto a los parámetros sin especificar (*"nuisance" parameters*)

El numerador es lo que se denomina un ***PROFILE LIKELIHOOD***

$$\lambda = \frac{L\left(\bar{x};\bar{\theta}_{r0},\hat{\hat{\theta}}_s\right)}{L\left(\bar{x};\hat{\theta}_r,\hat{\theta}_s\right)}$$

λ Es el cociente entre la "***profile likelihood***" y la "***maximum likelihood***"

En el denominador maximizamos la verosimilitud con respecto a todos los parámetros $\hat{\theta}_r,\hat{\theta}_s$

¡¡¡¡ Se ajusta dos veces !!!!

3.1 Test cociente de verosimilitud

Ejemplo: distribución normal

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0$$

σ : Sin especificar (*nuisance* parameter)

$$L\left(\overline{x};\mu,\sigma\right)=\left(2\pi\sigma^2\right)^{-n/2}\prod_{i=1}^{n}\exp\left[-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2\right]$$

Verosimilitud condicional (H0):

$$\left.\begin{aligned}
&\mu=\mu_0\\
&\hat{\hat{\sigma}}^2=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\mu_0\right)^2=\\
&\quad=\frac{1}{n}\left[n\left(\overline{x}-\mu_0\right)+\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2\right]=\\
&\quad=\left(\overline{x}-\mu_0\right)+s^2
\end{aligned}\right\}\Rightarrow$$

$$L\left(\overline{x};\mu_0,\hat{\hat{\sigma}}\right)=\left(2\pi\hat{\hat{\sigma}}^2\right)^{-n/2}\exp\left[-\frac{n}{2}\right]=$$

$$=\left[2\pi\left(\left(\overline{x}-\mu_0\right)+s^2\right)\right]^{-n/2}\exp\left[-\frac{n}{2}\right]$$

Verosimilitud incondicional:

La mejor estimación ML viene dada por:

$$\left.\begin{aligned}
&\hat{\mu}=\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i\\
&\hat{\sigma}^2=s^2=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2
\end{aligned}\right\}\Rightarrow$$

$$L\left(\overline{x};\hat{\mu},\hat{\sigma}\right)=\left(2\pi s^2\right)^{-n/2}\exp\left[-\frac{n}{2}\right]$$

3.1 Test cociente de verosimilitud

Ejemplo: distribución normal

Haciendo el cociente:

$$\lambda = \frac{L\left(\bar{x};\bar{\theta}_{r0},\hat{\hat{\theta}}_s\right)}{L\left(\bar{x};\hat{\bar{\theta}}_r,\hat{\bar{\theta}}_s\right)} = \frac{\left[2\pi\left(\left(\bar{x}-\mu_0\right)^2+s^2\right)\right]^{-n/2}\exp\left[-\frac{n}{2}\right]}{\left(2\pi s^2\right)^{-n/2}\exp\left[-\frac{n}{2}\right]} = \left\{\frac{s^2}{\left(\bar{x}-\mu_0\right)^2+s^2}\right\}^{n/2} = \left\{\frac{1}{1+\frac{t^2}{n-1}}\right\}^{n/2}$$

$$\lambda^{2/n} = \frac{1}{1+\frac{t^2}{n-1}} \quad \Rightarrow \quad t^2 = \frac{n\left(\bar{x}-\mu_0\right)^2}{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2} \quad \Rightarrow \quad t = \frac{\left(\bar{x}-\mu_0\right)}{s/\sqrt{n}}$$

t satisface una distribución t de Student con n-1 grados de libertad
(Tema 4 9.2)

Como además λ es una función monótona de t^2:

$$0 \le \lambda \le \lambda_\alpha \quad \Rightarrow \quad t^2 > t_\alpha^2 \quad \Rightarrow \quad -\infty < t < -t_{\alpha/2} \text{ y } t_{\alpha/2} < t < \infty$$

Double-sided test

$$\frac{\alpha}{2} = \int_{-\infty}^{-t_{\alpha/2}} f\left(t;n-1\right)dt = \int_{t_{\alpha/2}}^{\infty} f\left(t;n-1\right)dt$$

Si, $-t_{\alpha/2} < t_{\text{obs}} < t_{\alpha/2}$ aceptamos H_0

3.2 Distribución asintótica de λ

Para conocer la región crítica necesitamos conocer la pdf de λ asumiendo que H_0 es cierta. Posibilidades:

- Expresión analítica.
- Se suele recurrir al Monte Carlo, (***pseudo-experiments***).
- Considerar la distribución en el límite de grandes números (***distribución asintótica***).

En dicho límite la función de verosimilitud se comporta como una distribución normal:

Matriz de información

$$\bar{I} = \begin{bmatrix} \bar{I}_r & \vdots & \bar{I}_{rs} \\ \cdots & & \cdots \\ \bar{I}_{rs} & \vdots & \bar{I}_s \end{bmatrix}$$

$$L(\bar{x};\bar{\theta}) = L(\bar{x};\bar{\theta}_r,\bar{\theta}_s) \propto \exp\left[-\frac{1}{2}\left(\hat{\bar{\theta}}-\bar{\theta}\right)^T \bar{I}\left(\hat{\bar{\theta}}-\bar{\theta}\right)\right]$$

$$L(\bar{x};\bar{\theta}_r,\bar{\theta}_s) \propto \exp\left\{-\frac{1}{2}\left[\left(\hat{\bar{\theta}}_r-\bar{\theta}_r\right)^T \bar{I}_r\left(\hat{\bar{\theta}}_r-\bar{\theta}_r\right)+2\left(\hat{\bar{\theta}}_r-\bar{\theta}_r\right)^T \bar{I}_{rs}\left(\hat{\bar{\theta}}_s-\bar{\theta}_s\right)+\left(\hat{\bar{\theta}}_s-\bar{\theta}_s\right)^T \bar{I}_s\left(\hat{\bar{\theta}}_s-\bar{\theta}_s\right)\right]\right\}$$

En el máximo: $\hat{\bar{\theta}}_r = \bar{\theta}_r$ y $\hat{\bar{\theta}}_s = \bar{\theta}_s$ $\Rightarrow$ $L(\bar{x};\bar{\theta}) \propto 1$

Bajo la hipótesis de H_0: $\hat{\hat{\bar{\theta}}}_s = \bar{\theta}_s$ $\Rightarrow$ $L(\bar{x};\bar{\theta}) \propto \exp\left[-\frac{1}{2}\left(\bar{\theta}_r-\bar{\theta}_{0r}\right)^T \bar{I}_r\left(\bar{\theta}_r-\bar{\theta}_{0r}\right)\right]$

$$\lambda = \exp\left[-\frac{1}{2}\left(\bar{\theta}_r-\bar{\theta}_{0r}\right)^T \bar{I}_r\left(\bar{\theta}_r-\bar{\theta}_{0r}\right)\right] \Rightarrow -2\ln\lambda = \left(\bar{\theta}_r-\bar{\theta}_{0r}\right)^T \bar{I}_r\left(\bar{\theta}_r-\bar{\theta}_{0r}\right)$$

$$-2\ln\lambda \Rightarrow$$

En el límite asintótico -2lnλ se comporta como una distribución de χ^2

3.2 Distribución asintótica de λ

Bajo la hipótesis nula H0, ***-2lnλ*** se comporta como una distribución de $\chi^2(r)$ con *r* grados de libertad *(número de parámetros especificados)*

$$-2\ln\lambda = \left(\bar{\theta}_r - \bar{\theta}_{0r}\right)^T \bar{I}_r \left(\bar{\theta}_r - \bar{\theta}_{0r}\right)$$

Teorema de Wilk

Sea λ el cociente de verosimilitudes (***likelihood ratio***) para un test de hipótesis tal que:

$$H_0 : \bar{\theta} \in w \qquad \text{con } s \text{ parámetros libres (sin especificar)}$$

$$H_1 : \bar{\theta} \in \Omega - w \qquad \text{con } s + r \text{ parámetros libres (sin especificar)}$$

basado en una muestra de n variables aleatorias de la misma pdf. Bajo ciertas condiciones de regularidad y cuando la hipótesis nula H_0 es cierta, el test estadístico:

$-2\ln\lambda$ converge a una distribución de χ_r^2 en el límite $n \to \infty$

$$\lambda \to 0 \quad \Rightarrow \quad -2\ln\lambda \to \infty$$

Facilita el cálculo de la región crítica. Para un test de tamaño α, rechazamos H_0 si:

$$\lambda < \lambda_\alpha \qquad \Rightarrow \qquad -2\ln\lambda_\alpha > \chi_\alpha^2 \qquad\qquad \int_{\chi_\alpha^2}^{\infty} \chi^2(r)\, d\chi^2 = \alpha$$

3.3 Familias paramétricas diferentes

Cuando la pdf especificada de H_1 no se puede obtener variando los parámetros de la pdf de H_0, se habla de familias paramétricas diferentes:

$$H_0: \quad f\left(x:\bar{\phi}\right) \quad \text{con} \quad \phi \quad \text{sin especificar}$$
$$H_1: \quad g\left(x:\bar{\psi}\right) \quad \text{con} \quad \psi \quad \text{sin especificar}$$

Construimos una familia de funciones definiendo un nuevo parámetro θ

$$h\left(\bar{x}:\theta,\bar{\phi},\bar{\psi}\right)=(1-\theta)\, f\left(\bar{x}:\bar{\phi}\right)+\theta g\left(\bar{x}:\bar{\psi}\right)$$

Reformulamos testeando H_0 frete a H'_1

$$H_0: \quad h\left(\bar{x}:\theta,\bar{\phi},\bar{\psi}\right), \qquad \theta=0, \text{con } \bar{\phi},\bar{\psi} \text{ sin especificar}$$
$$H'_1: \quad h\left(\bar{x}:\theta,\bar{\phi},\bar{\psi}\right), \qquad \theta\neq 0, \text{con } \bar{\phi},\bar{\psi} \text{ sin especificar}$$

El cociente de verosimilitud vendrá dado por:

$$\lambda=\frac{L\left(\bar{x};\theta=0,\hat{\hat{\bar{\phi}}},\bar{\psi}\right)}{L\left(\bar{x};\hat{\theta},\hat{\bar{\phi}},\hat{\bar{\psi}}\right)}=\left[\frac{f\left(\bar{x};\hat{\hat{\bar{\phi}}}\right)}{(1-\theta)\, f\left(\bar{x}:\hat{\bar{\phi}}\right)+\theta g\left(\bar{x}:\hat{\bar{\psi}}\right)}\right]^n$$

Bajo la hipótesis nula H_0, ***-2lnλ*** se distribuye asintóticamente como una distribución de $\chi^2(1)$

(un solo parámetro especificado)

Bajo la hipótesis alternativa H_1, ***-2lnλ*** se distribuye asintóticamente como una distribución $\chi'^2(1,K)$ con 1 grado de libertad y parámetro central:

$$K=\frac{\theta^2}{S} \quad \text{con} \quad S=E\left[\frac{\left[f\left(\bar{x}:\bar{\phi}\right)-g\left(\bar{x}:\bar{\psi}\right)\right]^2}{\left[(1-\theta)\, f\left(\bar{x}:\bar{\phi}\right)+\theta g\left(\bar{x}:\bar{\psi}\right)\right]^2}\right]$$

3.4 Ejemplo. Búsquedas

Las hipótesis típicas a testear en caso de búsquedas son:

H_0 : "Solo Background".- *Todos los sucesos son de fondo"*

H_1 : "*Señal + Background*".- Los sucesos son una mezcla de los dos

Background

$$P(n|b) = \frac{b^n}{n!} e^{-b}$$

$$L_b = \frac{b^n}{n!} e^{-b} \prod_{i=1}^{n} f(x_i|b)$$

Señal + Background

$$P(n|s+b) = \frac{(s+b)^n}{n!} e^{-(s+b)}$$

$$L_{s+b} = \frac{(s+b)^n}{n!} e^{-(s+b)} \prod_{i=1}^{n} \left(\pi_s f(x_i|s) + \pi_b f(x_i|b) \right)$$

$$\pi_s = s/(s+b)$$
$$\pi_b = b/(s+b)$$

Definimos el test estadístico: $Q = -2\ln \frac{L_{s+b}}{L_b} = -2\left[-s + \sum_{i=1}^{n} \ln\left(1 + \frac{s}{b} \frac{f(x_i|s)}{f(x_i|b)} \right) \right]$

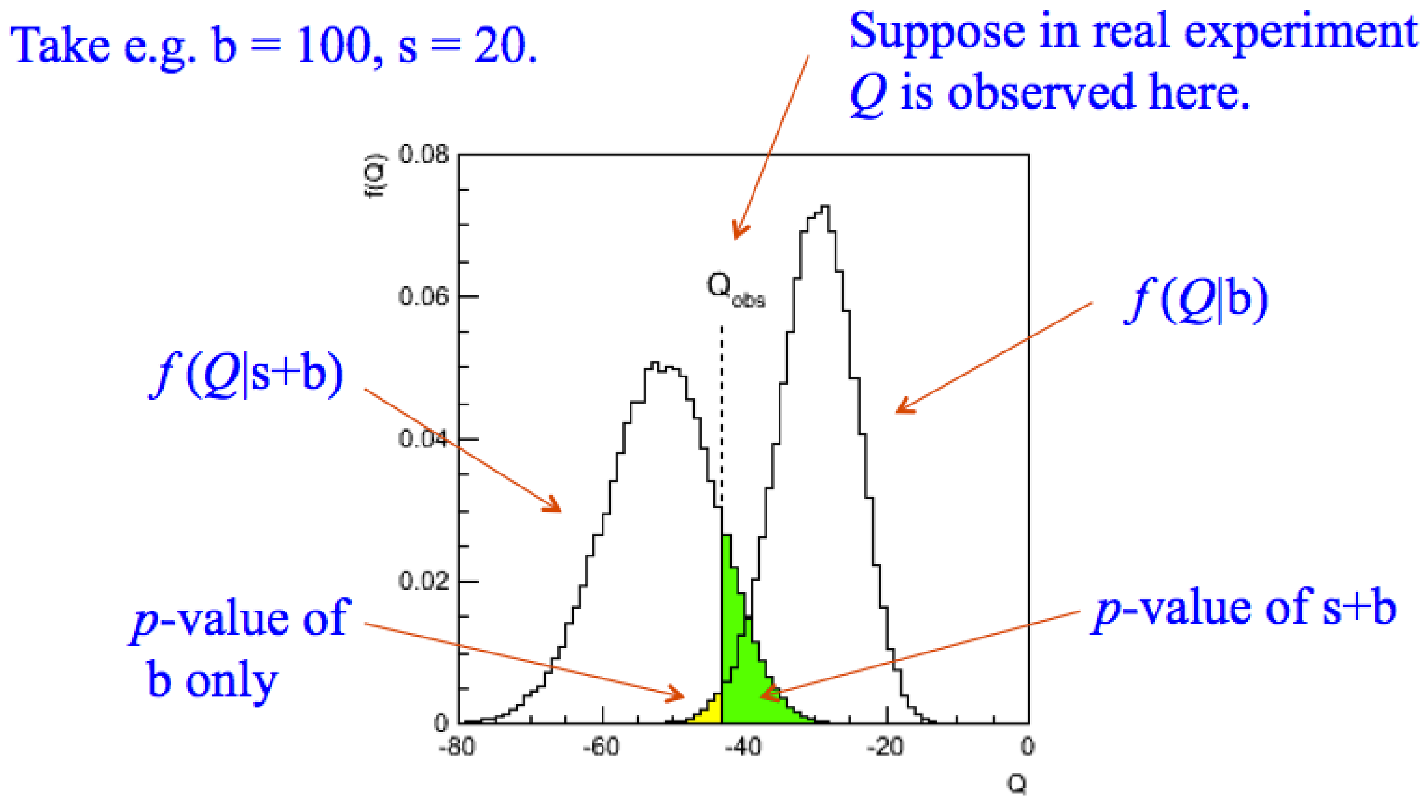
Para calcular el p-value necesitamos la pdf para Q bajo las diferentes hipótesis: $f(Q|b);\ f(Q|s+b)$

Utilizamos Monte Carlo toy (pseudoexperimentos).-

- Se genera muestras de pseudo-datos bajo la hipótesis de "**Solo background**" y por cada pseudo-experimento se calcula Q.
- Análogamente con la hipótesis "**Señal + fondo**".

3.4 Ejemplo. Búsquedas

Ejemplo



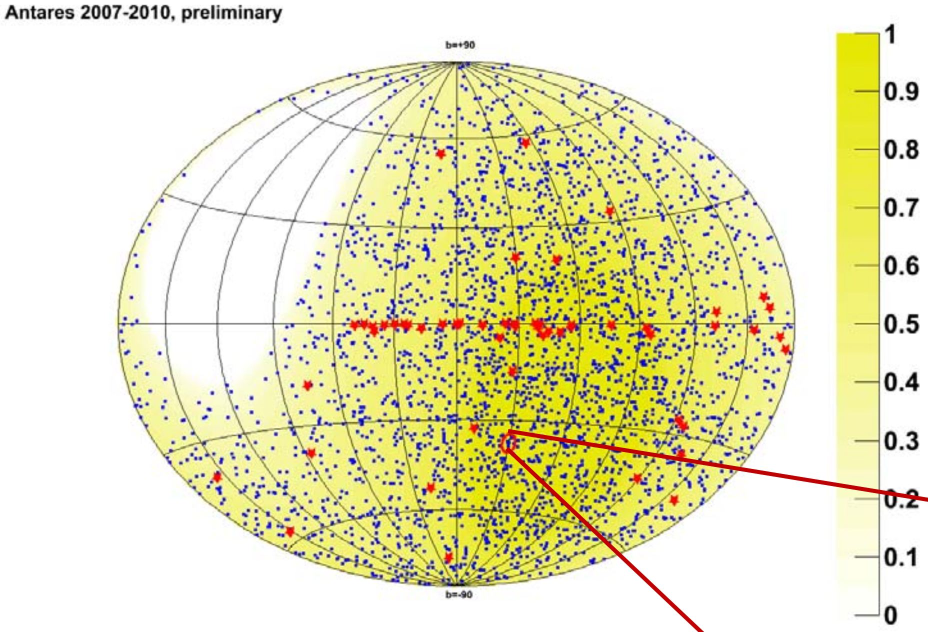
- **Amarillo** = p_b Si $p_b < 2.9\ 10^{-7}$ rechazamos la hipótesis de "**Solo background**" a 5 σ (discovery)
- **Verde** = p_{s+b} Si $p_{b+s} < \alpha$ rechazamos un modelo específico al 1- α de nivel de CL

3.4 Ejemplo. Búsquedas

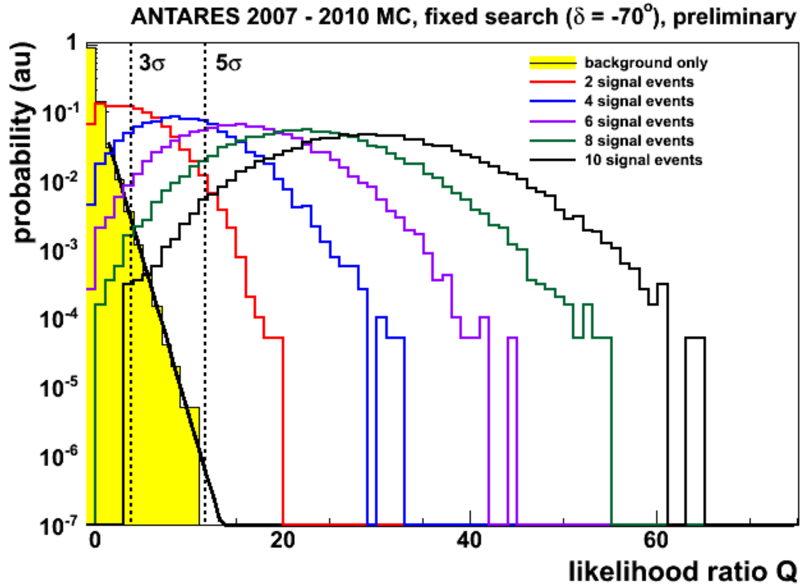
Ejemplo. Point-like neutrino sources

$$Q = L_{s+b}^{\max} - L_b$$

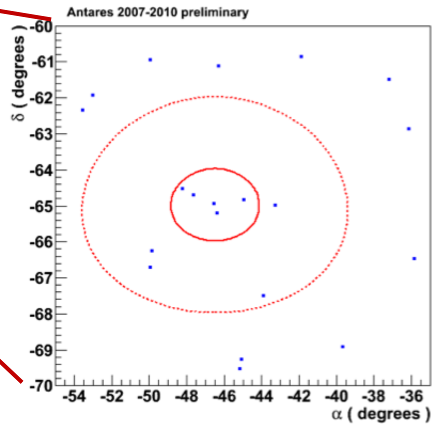
$$L_{s+b} = \sum_i \ln\left[\mu_s \mathcal{F}(\alpha_s, \delta_s) + \mathcal{B}(\delta_i)\right]$$



Skymap in galactic coordinates of the 3058 selected events. The red stars show the location of the 51 candidate sources considered in the fixed search and the red circle shows the location of the most signal like cluster found in the full sky search.



Distribution of the test statistic for the fixed search. The yellow histogram is for the background only case while the other lines are for different signal events added at a declination of -70 degrees.



Most significant cluster at:
$\mathcal{RA} = -46.5°$,
$\delta = -65.0°$

$N_{sig} = 5$ $Q = 13.02$
p-value = 0.026
Significance = 2.2 σ