

# PFVA'25 Kış Okulu Ders Notları

## **OLASILIK VE İSTATİSTİK II**

*(Parametre tahmini, Hipotez testleri,  
Poisson sayım, Güven aralığı)*

Orhan ÇAKIR

Ankara Üniversitesi, Fizik Bölümü  
[ocakir@science.ankara.edu.tr](mailto:ocakir@science.ankara.edu.tr)

# İçindekiler

- ❖ Giriş
- ❖ Hipotez, Olasılık
- ❖ Olasılık Fonksiyonu
- ❖ Parametre Tahmini
- ❖ Tahmin Ediciler
- ❖ Hipotez Testleri

# Giriş

**Parçacık fiziğinde** kullanılan istatistiksel yöntemlere genel bir bakış yapılmıştır. İstatistikte, olasılıkçı bir model hakkında çıkarımlar yapmak için belirli bir veri örneğini kullanmakla ilgileniyoruz, örneğin, modelin geçerliliğini değerlendirmek veya parametrelerinin değerlerini belirlemek için.

İstatistiksel çıkarıma yönelik iki ana yaklaşım vardır, bunlara Frekans yaklaşımı ve Bayes yaklaşımı diyebiliriz. **Frekans yaklaşımı istatistikte**, olasılık tekrarlanabilir bir deneyin sonucunun sınırlayıcı frekansı olarak yorumlanır. Bu çerçevedeki en önemli araçlar, parametre tahmini, istatistiksel testler ve belirli bir olasılıkla bir parametrenin gerçek değerini kapsayacak şekilde oluşturulan güven aralıklarıdır. Frekans yaklaşımı istatistikte, bir hipotez veya bir parametrenin değeri için önceden bir olasılık tanımlanmadığını vurgulayalım.

# Giriş

**Bayes yaklaşımı istatistikte**, olasılığın öznel yorumu, bir hipoteze olan kabul (inanma) derecesini ölçmek için kullanılır. Bu, bir parametrenin gerçek değerinin nerede bulunduğu dair bilginizi yansıtan bir olasılık yoğunluk fonksiyonunu (p.d.f.) tanımlamanıza olanak tanır. Bayes yöntemleri, genel olarak öznel olabilen ek bilgileri dahil etmek için doğal bir yol sağlar; aslında, ölçümü gerçekleştirmeden önce söz konusu hipotezler (veya parametreler) için önceki olasılıkları, yani parametrelerin değerleri hakkındaki kabul derecesini gerektirir. Bayes teoremini kullanarak, önceki kabul derecesi deneyden gelen verilerle güncellenir.

Birçok çıkarım problemi için, frekans ve Bayes yaklaşımları, farklı soruları yanıtlamalarına ve olasılığın temelde farklı yorumlarına dayanmalarına rağmen benzer sayısal değerler verir, Ref. [1]. Ancak bazı özel durumlarda, iki yaklaşım farklı sonuçlar verebilir.

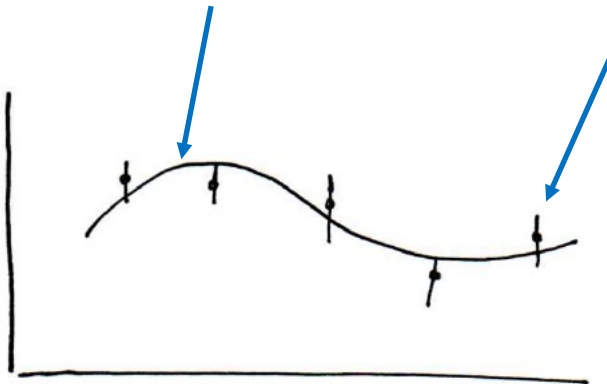
# Teori ↔ İstatistik ↔ Deney

Teori (model, hipotez):

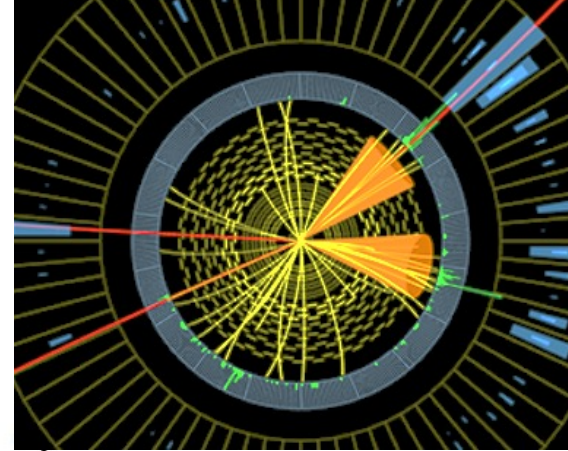
$$\begin{aligned}\mathcal{L} = & -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ & + i\bar{\psi}D\psi + h.c. \\ & + \bar{\psi}_i y_{ij} \psi_j \phi + h.c. \\ & + |D_\mu \phi|^2 - V(\phi)\end{aligned}$$

+ ölçüm aygıtının  
tepkisi

= model tahmini



Deney (gözlem):



Belirsizlik birçok  
seviyede dahil olur.

→ olasılık ile  
nicelenir!

# Hipotez, Olasılık

Bir deneyin tüm sonucunun (ölçüm kümesinin)  $x$  sayılarından oluşan bir koleksiyon olduğunu varsayalım.

(Basit) bir hipotez, her olası veri sonucuna bir olasılık atayan bir kuraldır:

$$P(\mathbf{x}|H) = H \text{ nin olasılığı}$$

Genellikle bir veya daha fazla belirsiz parametreyle etiketlenen bir hipotez ailesiyle uğraşırız (bir bileşik hipotez):

$$P(\mathbf{x}|\boldsymbol{\theta}) = L(\boldsymbol{\theta}) = \text{bu "olasılık fonksiyonu"}$$

↑ parametre

Not:

- 1) Olasılık için  $x$  verisini sabit olarak ele alıyoruz.
- 2) Olasılık fonksiyonu  $L(\theta)$ ,  $\theta$  için bir pdf değildir.

# Veri İçin Olasılık Fonksiyonu

\* Bağımsız ve özdeş dağılmış veri

Burada  $x$ 'in  $n$  bağımsız gözlemini ele alalım:  $x_1, \dots, x_n$ , burada  $x$ ,  $f(x; \theta)$ 'yi takip eder. Tüm veri örneği için ortak pdf şöyledir:

$$f(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta)$$

Bu durumda olasılık fonksiyonu

$$L(\vec{\theta}) = \prod_{i=1}^n f(x_i; \vec{\theta}) \quad (x_i \text{ sabit})$$

# Parametre Tahmini

Bir p.d.f' nin parametreleri onu karakterize eden sabitlerdir,

$$f(x; \theta) = \frac{1}{\theta} e^{-x/\theta}$$

burada,  $\theta$  hipotezlerin bir kümesini listeler

değişken      parametre

Gözlemlenen değerlerden oluşan bir örneğimiz olduğunu varsayalım :  $\mathbf{x} = (x_1, \dots, x_n)$

Parametreleri tahmin etmek için verilerin bazı işlevlerini bulmak istiyoruz:

$$\hat{\theta}(\vec{x})$$

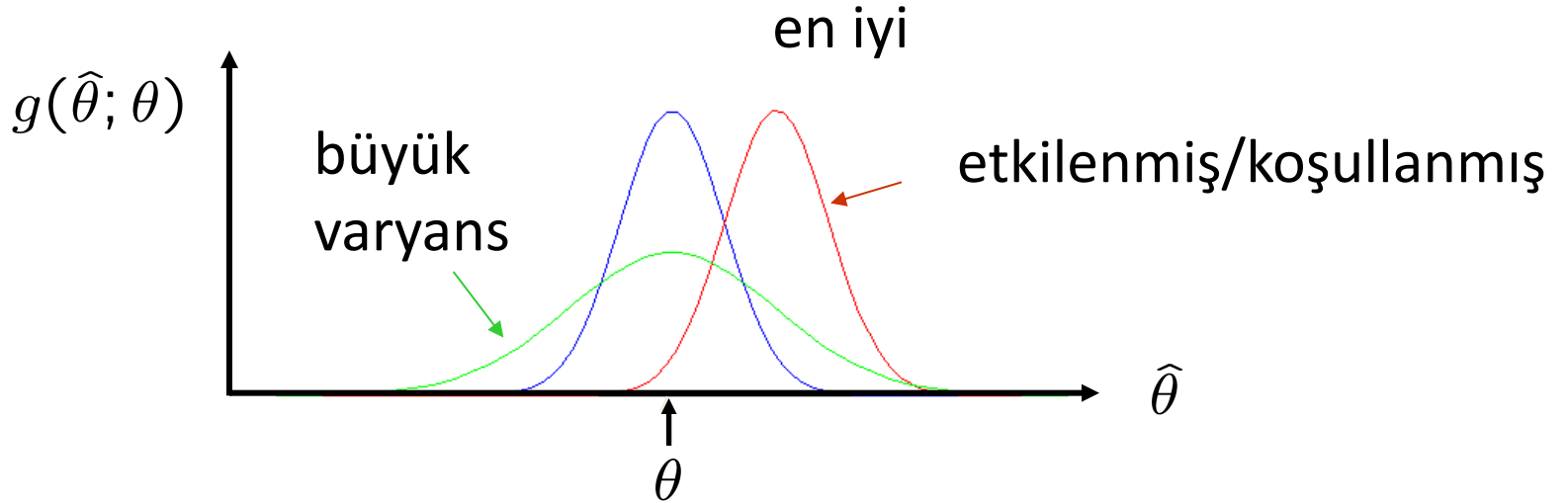
← ‘şapka’ ile yazılmış tahminci

Bazen  $x_1, \dots, x_n$  in bir fonksiyonu için ‘tahmin edici’ deriz; ‘tahmin’ belirli bir veri kümesindeki tahmin edicinin değeri içindir.



# Tahmin Edicilerin Özellikleri

Tüm ölçümü tekrarlayacak olsaydık, her birinden gelen tahminler bir pdf'yi takip ederdi:



Küçük (veya sıfır) önyargı (sistemik hata) istiyoruz:  $b = E[\hat{\theta}] - \theta$

→ tekrarlanan ölçümlerin ortalaması gerçek değere yakın olmalıdır.

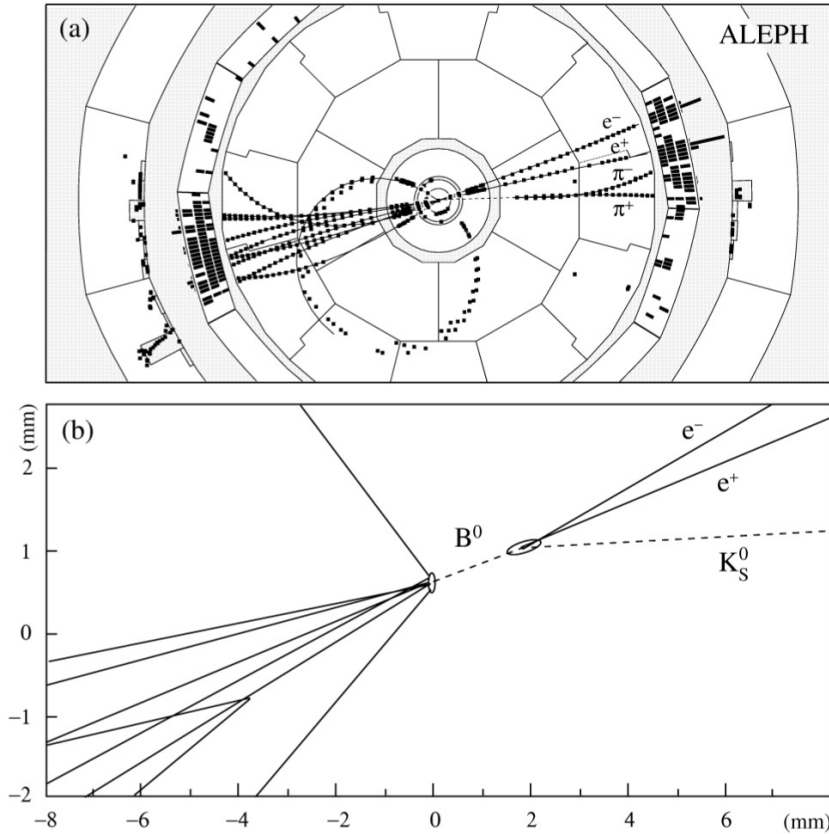
ve küçük bir varyans (istatistiksel hata) istiyoruz:  $V[\hat{\theta}]$

→ küçük önyargı & varyans genel olarak çelişkili kriterlerdir.

# Örnek: kararsız bir parçacığın bozunumu

Birkaç istatistiksel yöntemi açıklamak için kullanacağımız bir örnek olarak, B mezonu gibi kararsız bir parçacığın uygun bozunma süresini ölçmeyi düşünelim:

R. Barate et al. / Physics Letters B 492 (2000) 259–274



Kütlesi  $m_B$  olan  $B^0$  mezounun bozunma ürünlerinin uçuş mesafesi  $d$  ve momentum  $p$ 'yi ölçün.

Bunlar, uygun (proper) bozunma süresi  $t_p$  ( $B^0$  durağan çerçevesindeki zaman) ile ilişkilidir.

$$d = vt_{\text{lab}} = \beta c \times \gamma t_p = \frac{p_B}{m_B} t_p$$

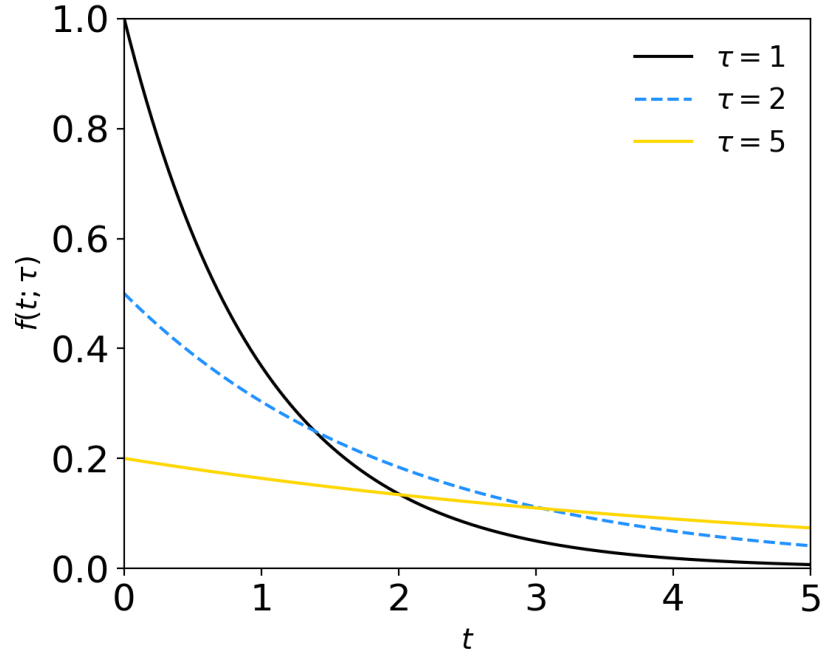
$$t_p = \frac{m_B d}{p_B}$$

# Uygun bozunma süresi için üstel pdf

Zaman  $t$ 'yi aşağıdaki gibi üstel bir pdf olarak modelleyebiliriz:

$$f(t; \tau) = \frac{1}{\tau} e^{-t/\tau}, \quad t \geq 0$$

↑ rastgele değişken      ↑ parametre



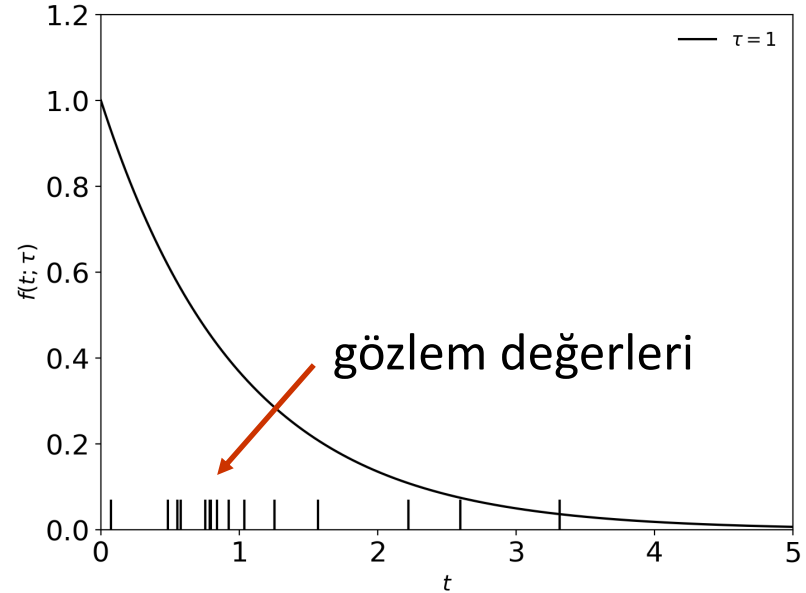
Bu  $t$ 'nin ortalaması ve varyansı şu şekilde olduğunu gösterebiliriz:

$$E[t] = \int_0^{\infty} t f(t; \tau) dt = \tau$$

$$V[t] = E[t^2] - (E[t])^2 = \tau^2$$
$$\sigma = \sqrt{V(t)} = \tau$$

# Ödev: Üstel olarak dağıtılmış verilerle istatistik

Haftanın ilerleyen günlerinde;  
Deneyin  $n$  kez tekrarlandığını  
ve şu verileri verdiğini  
varsayalım:  $t_1, \dots, t_n$ .



Veri değerlerini kullanarak, ortalama ömür  $\tau$ 'yi tahmin edin.

Tahmin içindeki istatistiksel belirsizliği niceleyin.

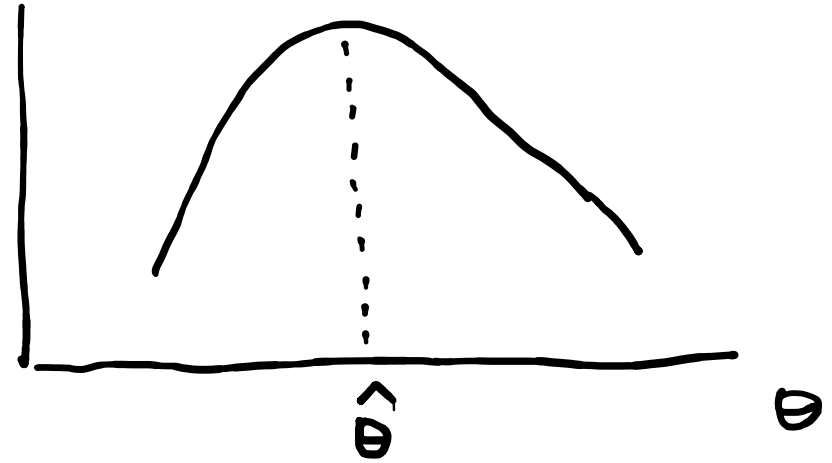
Ortalama ömür için üst/alt sınırları bildirin.

# Maksimum Olasılık Tahmin Ediciler (MLE)

Maksimum olasılık tahmin edicilerini veya MLE'leri, olasılığın maksimum olduğu parametre değerleri olarak tanımlıyoruz.

L'yi maksimize etmek,  
log L'yi maksimize  
etmeye eşdeğerdir

$\ln L$



$$\hat{\theta} = \operatorname{argmax}_{\theta} L(\theta)$$

Birden fazla maksimuma sahip olabilir (en yüksek olanı al).

MLE'lerin herhangi bir 'optimum' özelliğe sahip olması garanti edilmez (ancak pratikte çok iyidir).

# MLE Örneği: üstel pdf parametresi

Üstel pdf'yi düşünün,  $f(t; \tau) = \frac{1}{\tau} e^{-t/\tau}$

ve diyelim ki verilerimiz var,  $t_1, \dots, t_n$

Olasılık fonksiyonları  $L(\tau) = \prod_{i=1}^n \frac{1}{\tau} e^{-t_i/\tau}$

$L(\tau)$ 'nin maksimum olduğu  $\tau$  değeri aynı zamanda logaritmasının (log olasılık fonksiyonunun) maksimum değerini de verir:

$$\ln L(\tau) = \sum_{i=1}^n \ln f(t_i; \tau) = \sum_{i=1}^n \left( \ln \frac{1}{\tau} - \frac{t_i}{\tau} \right)$$

# MLE Örneği: üstel pdf parametresi (2)

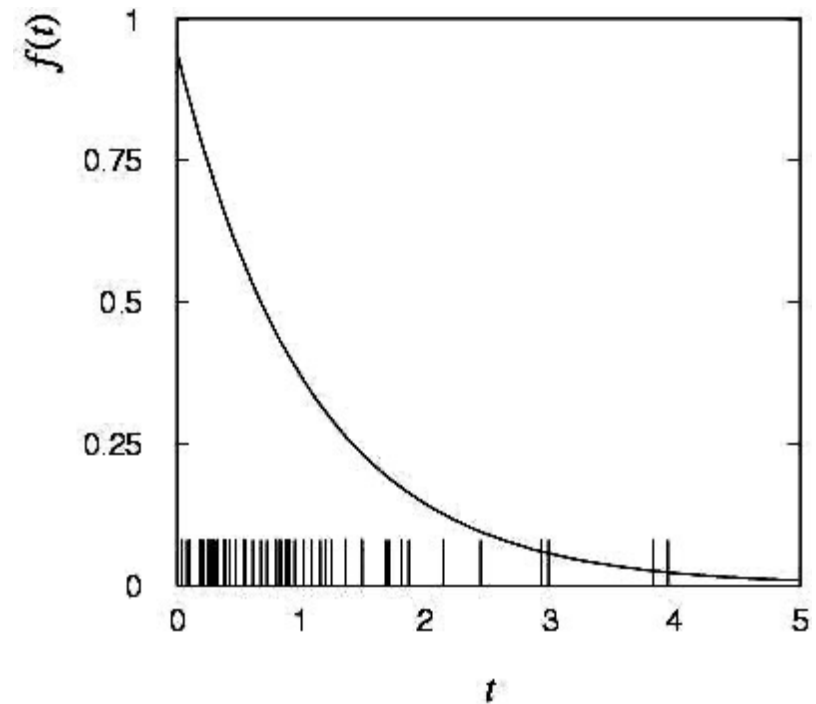
Maksimum değerini bulun  $\frac{\partial \ln L(\tau)}{\partial \tau} = 0$ ,

$$\rightarrow \hat{\tau} = \frac{1}{n} \sum_{i=1}^n t_i$$

Monte Carlo test için:  
50 değer üretin  
 $\tau = 1$  kullanın:

ML tahmin ediciyi buluruz:

$$\hat{\tau} = 1.062$$



# MLE Örneği: üstel pdf parametresi (3)

Üstel dağılım için ortalama ve varyans:

$$E[t] = \int_0^{\infty} t \frac{1}{\tau} e^{-t/\tau} dt = \tau$$

$$V[t] = \int_0^{\infty} (t - \tau)^2 \frac{1}{\tau} e^{-t/\tau} dt = \tau^2$$

MLE için bu  $\hat{\tau} = \frac{1}{n} \sum_{i=1}^n t_i$  ve aşağıdakileri elde ederiz:

$$E[\hat{\tau}] = E \left[ \frac{1}{n} \sum_{i=1}^n t_i \right] = \frac{1}{n} \sum_{i=1}^n E[t_i] = \tau \quad \longrightarrow \quad b = E[\hat{\tau}] - \tau = 0$$

$$V[\hat{\tau}] = V \left[ \frac{1}{n} \sum_{i=1}^n t_i \right] = \frac{1}{n^2} \sum_{i=1}^n V[t_i] = \frac{\tau^2}{n} \quad \longrightarrow \quad \sigma_{\hat{\tau}} = \frac{\tau}{\sqrt{n}}$$



# Tahmin Ediciler Varyansı: MC yöntemi

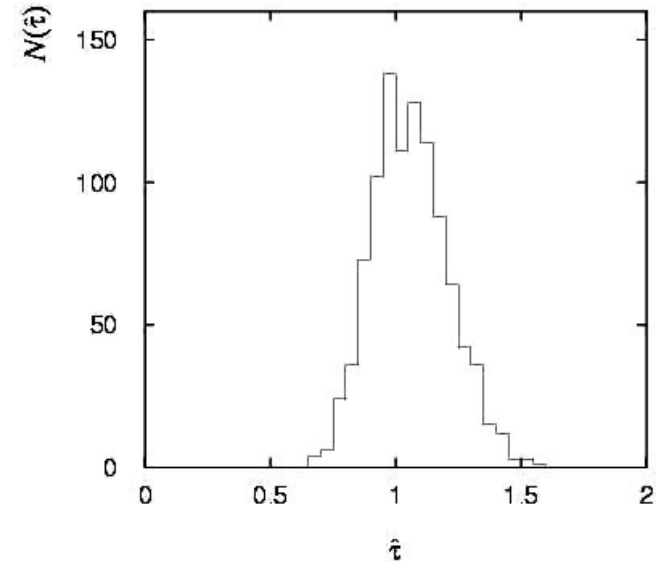
Parametremizi tahmin ettikten sonra şimdi onun 'istatistiksel hatasını', yani tüm ölçümü birçok kez tekrarlırsak tahminlerin ne kadar yaygın olarak dağılacığını bildirmemiz gerekiyor.

Bunu yapmanın bir yolu, tüm deneyi bir Monte Carlo programıyla birçok kez simüle etmek olacaktır (MC için ML tahmini kullanın).

Üstel örnek için, tahminlerin örneklem varyansından şunu buluruz:

$$\hat{\sigma}_{\hat{\tau}} = 0.151$$

Tahminlerin dağılımının yaklaşık olarak Gauss olduğunu unutmayın – büyük örneklem limitinde ML için (neredeyse) her zaman doğrudur.



# Tahmin Edicilerin Varyansı (RCF'den gelen)

Bilgi eşitsizliği / Rao-Cramer-Frechet (RCF) eşitsizliği, herhangi bir tahmin edicinin (sadece ML değil) varyansına bir alt sınır koyar:

$$V[\hat{\theta}] \geq \left(1 + \frac{\partial b}{\partial \theta}\right)^2 / E \left[ -\frac{\partial^2 \ln L}{\partial \theta^2} \right]$$

Minimum Varyans Sınırı (MVB)  
( $b = E[\hat{\theta}] - \theta$ )

Genellikle  $b$  önyargısı küçüktür ve eşitlik ya tam olarak geçerlidir ya da iyi bir yaklaşımdır (örn. büyük veri örneği sınırı). Sonra,

$$V[\hat{\theta}] \approx -1 / E \left[ \frac{\partial^2 \ln L}{\partial \theta^2} \right]$$

Bunu,  $\ln L$ 'nin maksimum noktasındaki 2. türevini kullanarak tahmin edin:

$$\hat{V}[\hat{\theta}] = - \left( \frac{\partial^2 \ln L}{\partial \theta^2} \right)^{-1} \Bigg|_{\theta=\hat{\theta}}$$

# Üstel parametrelili MLE için MVB

işlem: 
$$\text{MVB} = - \left( 1 + \frac{\partial b}{\partial \tau} \right)^2 / E \left[ \frac{\partial^2 \ln L}{\partial \tau^2} \right]$$

Üstel parametre için MLE'yi buluruz

$$\hat{\tau} = \frac{1}{n} \sum_{i=1}^n t_i$$

ve  $b = 0$  olduğunu gösteririz, böylece  $\partial b / \partial \tau = 0$ .

işlem: 
$$\frac{\partial^2 \ln L}{\partial \tau^2} = \sum_{i=1}^n \left( \frac{1}{\tau^2} - \frac{2t_i}{\tau^3} \right)$$

ve bütün  $i$  lerde  $E[t_i] = \tau$  dir 
$$E \left[ \frac{\partial^2 \ln L}{\partial \tau^2} \right] = -\frac{n}{\tau^2},$$

ve böylece 
$$\text{MVB} = \frac{\tau^2}{n} = V[\hat{\tau}]. \quad (\text{Burada MLE "verimli"}).$$

# Tahmin edicilerin varyansı: grafiksel yöntem

Burada  $\ln L(\vartheta)$  yı maksimumu civarında açarız:

$$\ln L(\theta) = \ln L(\hat{\theta}) + \left[ \frac{\partial \ln L}{\partial \theta} \right]_{\theta=\hat{\theta}} (\theta - \hat{\theta}) + \frac{1}{2!} \left[ \frac{\partial^2 \ln L}{\partial \theta^2} \right]_{\theta=\hat{\theta}} (\theta - \hat{\theta})^2 + \dots$$

İlk terim  $\ln L_{\max}$ , ikinci terim sıfır, ve üçüncü terim için bilgi eşitsizliği (eşitlik varsayılır) kullanılır:

$$\ln L(\theta) \approx \ln L_{\max} - \frac{(\theta - \hat{\theta})^2}{2\widehat{\sigma}_{\hat{\theta}}^2}$$

ifadesi,  $\ln L(\hat{\theta} \pm \hat{\sigma}_{\hat{\theta}}) \approx \ln L_{\max} - \frac{1}{2}$

→  $\hat{\sigma}_{\hat{\theta}}$  alınır,  $\theta$  değiştirilir,

$\hat{\theta}$  dan itibaren,  $\ln L$  değeri  $\frac{1}{2}$  faktör azalır.

# Grafiksel yöntemle varyans örneği

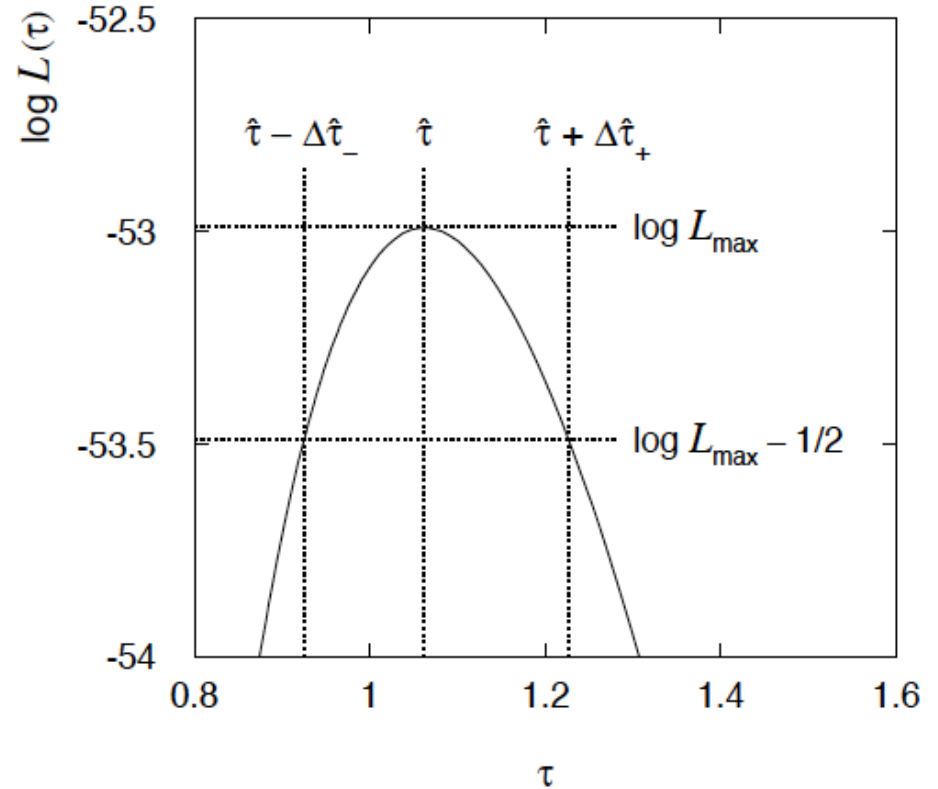
Üstel ile ML örneği:

$$\hat{\tau} = 1.062$$

$$\Delta\hat{\tau}_- = 0.137$$

$$\Delta\hat{\tau}_+ = 0.165$$

$$\hat{\sigma}_{\hat{\tau}} \approx \Delta\hat{\tau}_- \approx \Delta\hat{\tau}_+ \approx 0.15$$



Sonlu örneklem büyüklüğü ( $n = 50$ ) olduğundan  $\ln L$ 'de tam olarak parabolik değil.

# N parametre için bilgi eşitsizliği

N parametrenin tahmin edildiğini varsayalım  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_r)$

*Fisher bilgi matrisi*

$$I_{ij} = -E \left[ \frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \right] = - \int \frac{\partial^2 \ln P(\mathbf{x}|\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} P(\mathbf{x}|\boldsymbol{\theta}) d\mathbf{x}$$

ve tahmin edicilerin kovaryans matrisi :  $\hat{\boldsymbol{\theta}} \quad V_{ij} = \text{cov}[\hat{\theta}_i, \hat{\theta}_j]$

Bilgi eşitsizliği matrisin şu şekilde olduğunu belirtir:

$$M_{ij} = V_{ij} - \sum_{k,l} \left( \delta_{ik} + \frac{\partial b_i}{\partial \theta_k} \right) I_{kl}^{-1} \left( \delta_{lj} + \frac{\partial b_l}{\partial \theta_j} \right)$$

pozitif yarı tanımlı / kesin:

$$\mathbf{z}^T M \mathbf{z} \geq 0 \text{ bütün } \mathbf{z} \neq 0, \text{ diyagonal elemanlar } \geq 0$$

## N parametre için bilgi eşitsizliği (2)

Pratikte eşitsizlik ~ her zaman büyük örneklem sınırında kullanılır:

bias  $\rightarrow 0$

eşitsizlik  $\rightarrow$  eşitlik, yani  $M = 0$ , ve böylece  $V^{-1} = I$

Burada,  $V_{ij}^{-1} = -E \left[ \frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \right]$

Bu ifade ile data' dan tahmin edilebilir:  $\hat{V}_{ij}^{-1} = - \frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \Big|_{\hat{\theta}}$

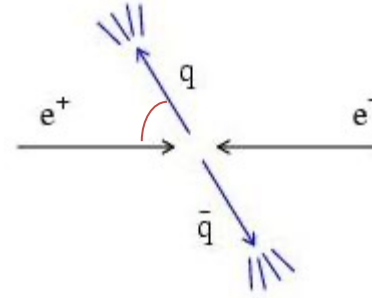
$V^{-1}$  matrisini sayısal olarak (veya otomatik türevleme ile) bulun, ardından tahmin edicilerin kovaryans matrisini elde etmek için tersini alın

$$\hat{V}_{ij} = \widehat{\text{cov}}[\hat{\theta}_i, \hat{\theta}_j]$$

# İki parametrelili ML örneđi

Örnek: Bir saçılma açısı dağılımını göz önünde bulundurun,  $x = \cos \theta$ ,

$$f(x; \alpha, \beta) = \frac{1 + \alpha x + \beta x^2}{2 + 2\beta/3}$$



Deđişken deđer aralıđı  $x_{\min} < x < x_{\max}$ , normalize edilmesi gerekir:

$$\int_{x_{\min}}^{x_{\max}} f(x; \alpha, \beta) dx = 1 .$$

Örnek:  $\alpha = 0.5$ ,  $\beta = 0.5$ ,  $x_{\min} = -0.95$ ,  $x_{\max} = 0.95$ , burada MC ile  $n = 2000$  olay üretin.

$$\ln L(\alpha, \beta) = \sum_{i=1}^n \ln f(x_i; \alpha, \beta)$$

← Maksimumu sayısal olarak buluruz



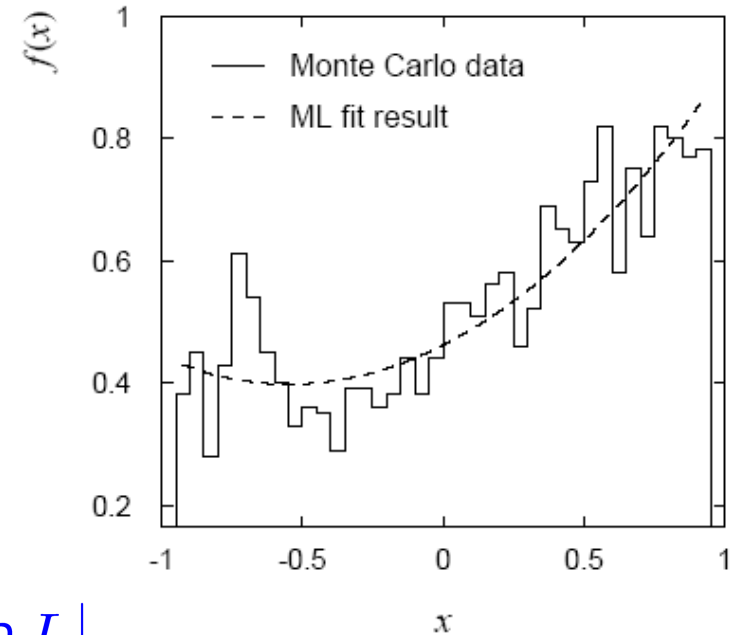
# iki parametrelili ML örneđi: fit sonuçları

$\ln L(\alpha, \beta)$  'nın maksimumunu sayısal olarak bulunur:

$$\hat{\alpha} = 0.508$$

$$\hat{\beta} = 0.47$$

Not: Uyum için verilerin gruplara ayrılmasına gerek yoktur, ancak uyum iyiliđi için histogramla karşılaştırma yapılabilir (örn. 'görsel' veya  $\chi^2$ ).



(Ko)varyanslar  $(\widehat{V}^{-1})_{ij} = -\frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \Big|_{\vec{\theta} = \hat{\vec{\theta}}}$

$$\hat{\sigma}_{\hat{\alpha}} = 0.052$$

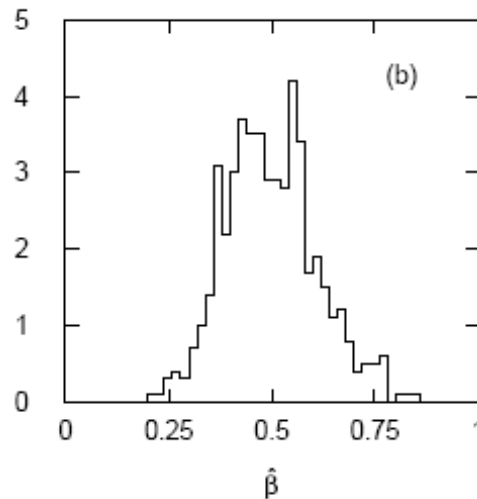
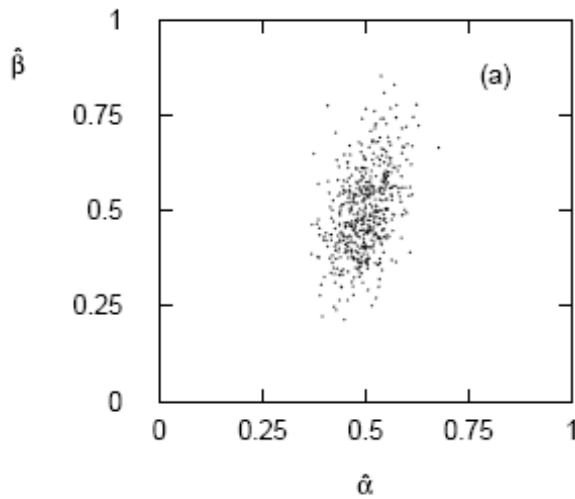
$$\text{cov}[\hat{\alpha}, \hat{\beta}] = 0.0026$$

$$\hat{\sigma}_{\hat{\beta}} = 0.11$$

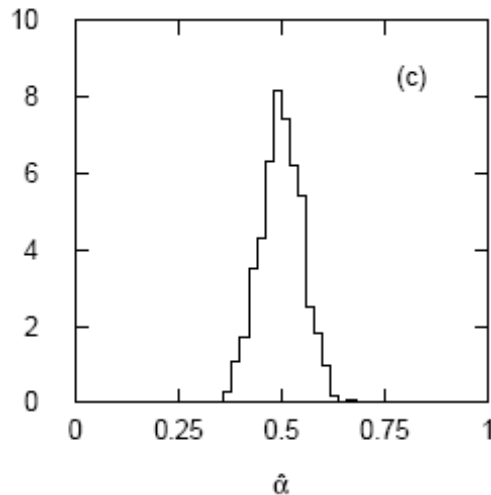
$$r = 0.46 = \text{korelasyon katsayısı}$$

# İki-parametrelili fit: MC çalışması

500 deneme, hepsi  $n = 2000$  olayla ML uyumunu tekrarlayın:



$$\begin{aligned}\bar{\hat{\alpha}} &= 0.499 \\ s_{\hat{\alpha}} &= 0.051 \\ \bar{\hat{\beta}} &= 0.498 \\ s_{\hat{\beta}} &= 0.111 \\ \widehat{\text{cov}}[\hat{\alpha}, \hat{\beta}] &= 0.0024 \\ r &= 0.42\end{aligned}$$



Tahminler  $\sim$ gerçek değerlere ortalama;  
(K)ovaryanslar önceki tahminlere yakın;  
marjinal pdf'ler yaklaşık olarak Gauss'tur.

# Varyanslar için çok parametreliliksel yöntem

$\ln L(\boldsymbol{\theta})$  'nın açılımını yapılır, MLE etrafında 2<sup>nci</sup> mertebeden:

$$\ln L(\boldsymbol{\theta}) \approx \ln L(\hat{\boldsymbol{\theta}}) + \sum_i \frac{\partial \ln L}{\partial \theta_i} \Big|_{\hat{\boldsymbol{\theta}}} (\theta_i - \hat{\theta}_i) + \frac{1}{2!} \sum_{i,j} \frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \Big|_{\hat{\boldsymbol{\theta}}} (\theta_i - \hat{\theta}_i)(\theta_j - \hat{\theta}_j)$$

$\ln L_{\max}$

sıfır

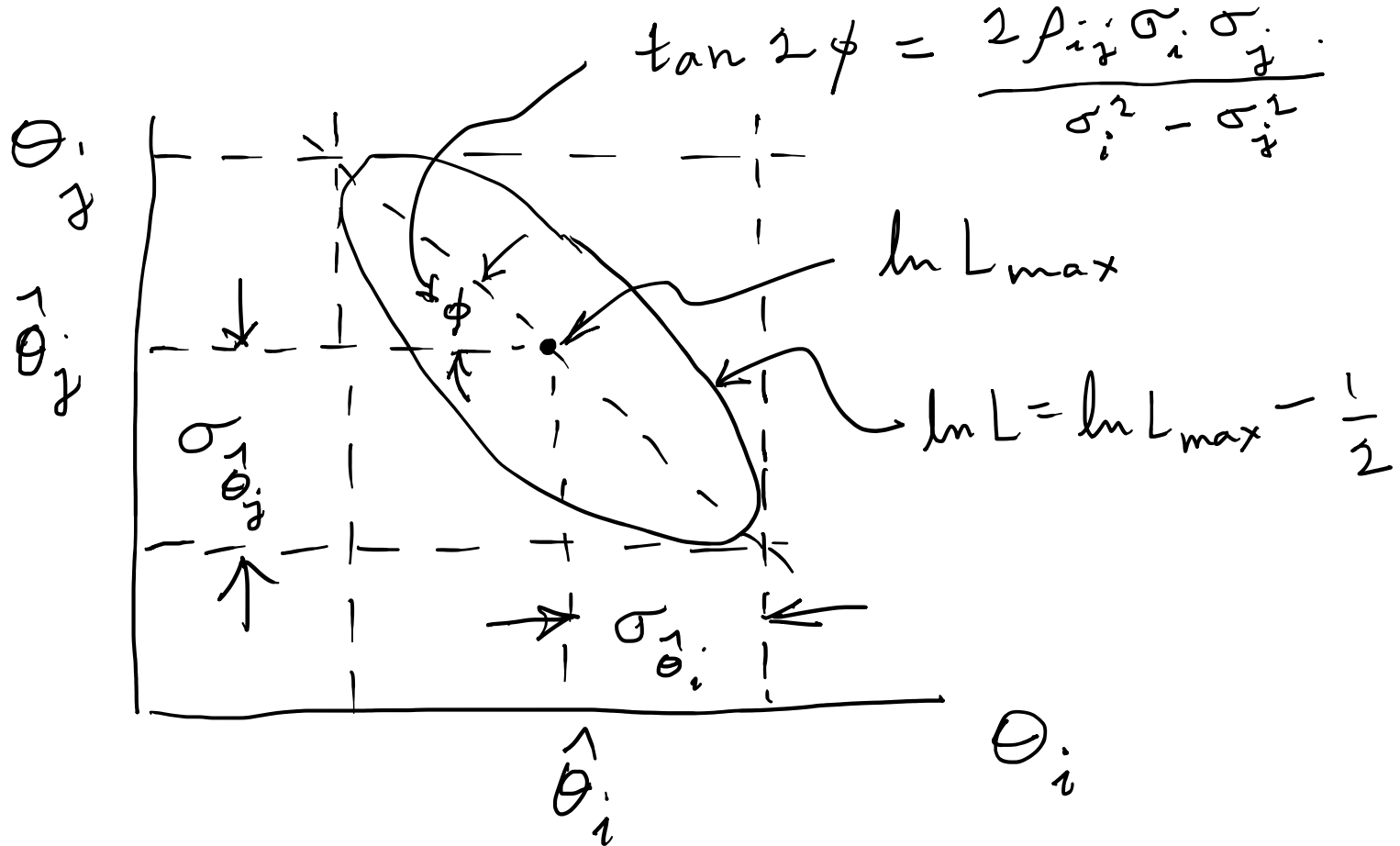
Bilgi eşit(siz)liği kullanılarak MLE'lerin kovaryans matrisiyle ilişkilendirilir.

**Sonuç:**  $\ln L(\boldsymbol{\theta}) = \ln L_{\max} - \frac{1}{2} \sum_{i,j} (\theta_i - \hat{\theta}_i) V_{ij}^{-1} (\theta_j - \hat{\theta}_j)$

Böylece yüzey  $\ln L(\boldsymbol{\theta}) = \ln L_{\max} - \frac{1}{2}$  aşağıdakine karşı gelir

$(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T V^{-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) = 1$  , bu ise bir (hiper-) elips denklemdir.

## Çok parametrelı grafıksel yöntem (2)



MLE'den tanjant düzlemlere olan uzaklık standart sapmaları verir.

# İki parametre için $\ln L_{\max} - 1/2$ konturu

Büyük  $n$  için,  $\ln L$  maksimum yakınında ikinci dereceden bir form alır:

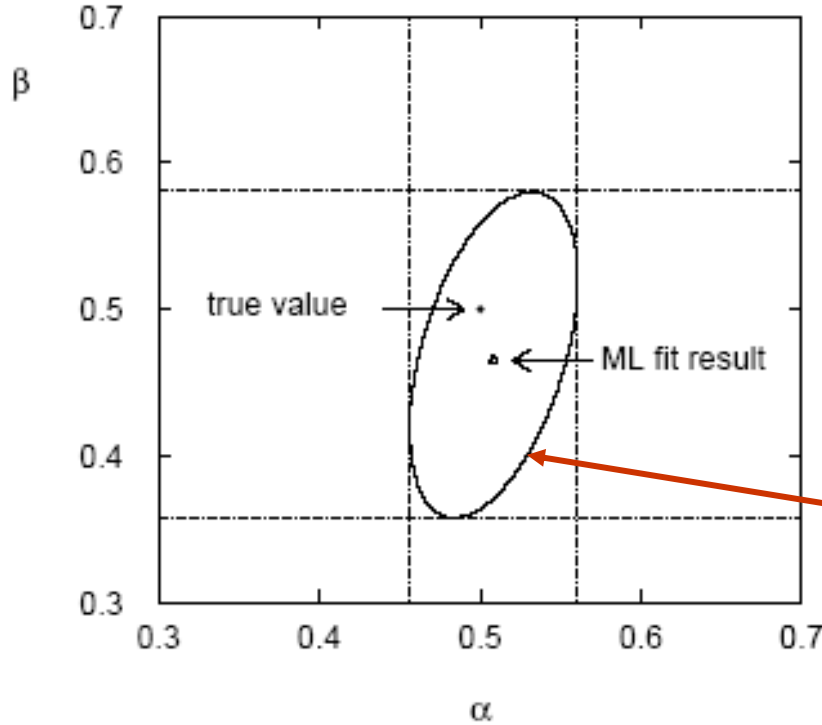
$$\ln L(\alpha, \beta) \approx \ln L_{\max}$$

$$-\frac{1}{2(1 - \rho^2)} \left[ \left( \frac{\alpha - \hat{\alpha}}{\sigma_{\hat{\alpha}}} \right)^2 + \left( \frac{\beta - \hat{\beta}}{\sigma_{\hat{\beta}}} \right)^2 - 2\rho \left( \frac{\alpha - \hat{\alpha}}{\sigma_{\hat{\alpha}}} \right) \left( \frac{\beta - \hat{\beta}}{\sigma_{\hat{\beta}}} \right) \right]$$

Kontur  $\ln L(\alpha, \beta) = \ln L_{\max} - 1/2$  bir elipstir:

$$\frac{1}{(1 - \rho^2)} \left[ \left( \frac{\alpha - \hat{\alpha}}{\sigma_{\hat{\alpha}}} \right)^2 + \left( \frac{\beta - \hat{\beta}}{\sigma_{\hat{\beta}}} \right)^2 - 2\rho \left( \frac{\alpha - \hat{\alpha}}{\sigma_{\hat{\alpha}}} \right) \left( \frac{\beta - \hat{\beta}}{\sigma_{\hat{\beta}}} \right) \right] = 1$$

# In L konturundan (k)ovaryanslar



MC verisi için  
 $\alpha, \beta$  düzlemi

$$\ln L(\alpha, \beta) = \ln L_{\max} - 1/2$$

→ Konturlara tanjant çizgileri standard sapmaları verir.

→ Elips açıları  $\phi$ , korelasyonla ilişkilidir:  $\tan 2\phi = \frac{2\rho\sigma_{\hat{\alpha}}\sigma_{\hat{\beta}}}{\sigma_{\hat{\alpha}}^2 - \sigma_{\hat{\beta}}^2}$

# Frekansçı (Frequentist) hipotez testleri

Bir ölçümün  $x$  verisi ürettiğini varsayalım; test etmek istediğimiz bir  $H_0$  hipotezi ve alternatif bir  $H_1$  hipotezi düşünün

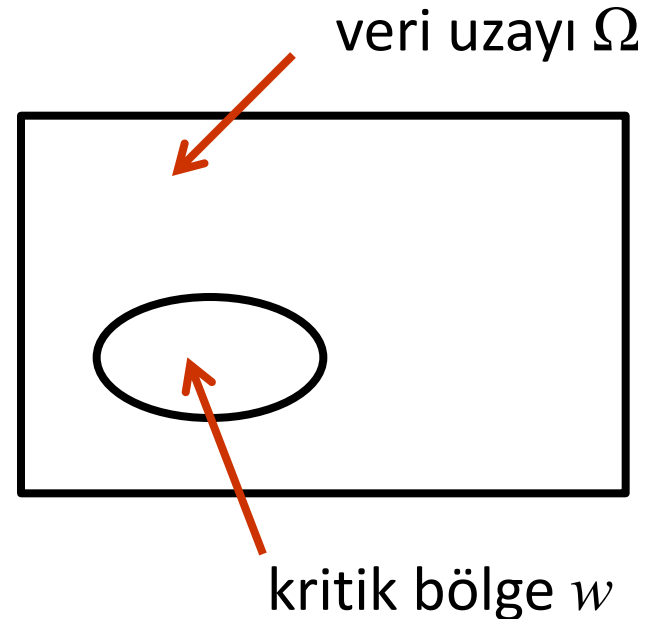
$H_0, H_1$  burada  $x$  için olasılık belirtir:  $P(x|H_0), P(x|H_1)$

$H_0$  in testi, veri alanının  $w$  kritik bölgesini belirleyerek tanımlanır, böylece  $H_0$ 'ın doğru olduğu varsayıldığında, oradaki verileri gözlemlemek için (küçük) bir olasılık  $\alpha$ 'dan daha fazla bir olasılık olmaz, yani,  $P(x \in w | H_0) \leq \alpha$

veri kesikli ise, eşitsizlik uygulanır.

Burada  $\alpha$  testin önem seviyesi veya büyüklüğü olarak isimlendirilir.

Eğer  $x$  kritik bölgede gözlenirse,  $H_0$  red edilir.

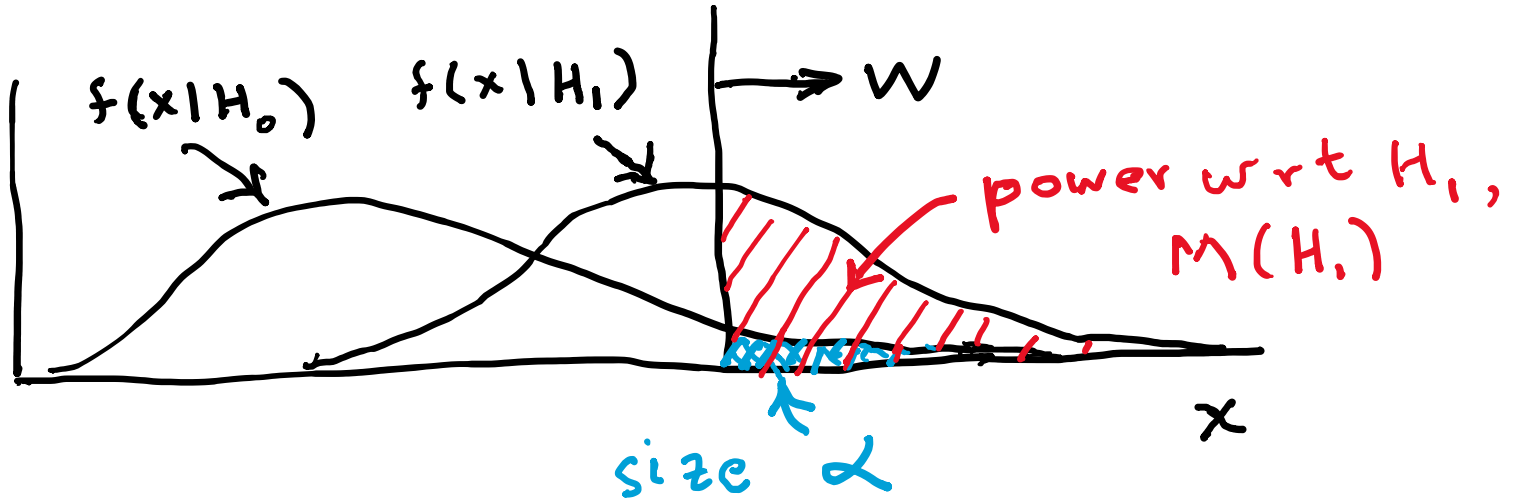


## Bir testin tanımlanması (2)

Genel olarak aynı  $\alpha$  boyutunu veren sonsuz sayıda olası kritik bölge vardır.

Kritik bölgenin nereye yerleştirileceğini motive etmek için alternatif hipotez  $H_1$ 'i kullanın.

Kabaca söylemek gerekirse,  $H_0$  doğruysa düşük bir olasılık ( $\alpha$ )'nın bulunduğu, ancak  $H_1$  doğruysa yüksek bir olasılığın bulunduğu kritik bölgeyi yerleştirin:





# Sınıflandırma: istatistiksel bir test

Olayların iki olası türde olduğunu varsayalım:

s (sinyal) ve b (ardalan / backgr.)

Her olay için, arka plan olması hipotezini test edin, yani  $H_0 = b$ .

Her biri s veya b türünde olan birçok olay üzerinde test yapın, yani burada hipotez, olaydan olaya rastgele değişen "gerçek sınıf etiketi"dir, bu nedenle buna bir frekansçı olasılık atayabiliriz.

$H_0$ 'ın "s türünde aday olaylar" olarak reddedildiği olayları seçin, eşdeğer Parçacık Fiziği terminolojisi:

ardalan verimliliği  $\epsilon_b = \int_W f(\mathbf{x}|H_0) d\mathbf{x} = \alpha$

sinyal verimliliği:  $\epsilon_s = \int_W f(\mathbf{x}|H_1) d\mathbf{x} = 1 - \beta = \text{power}$

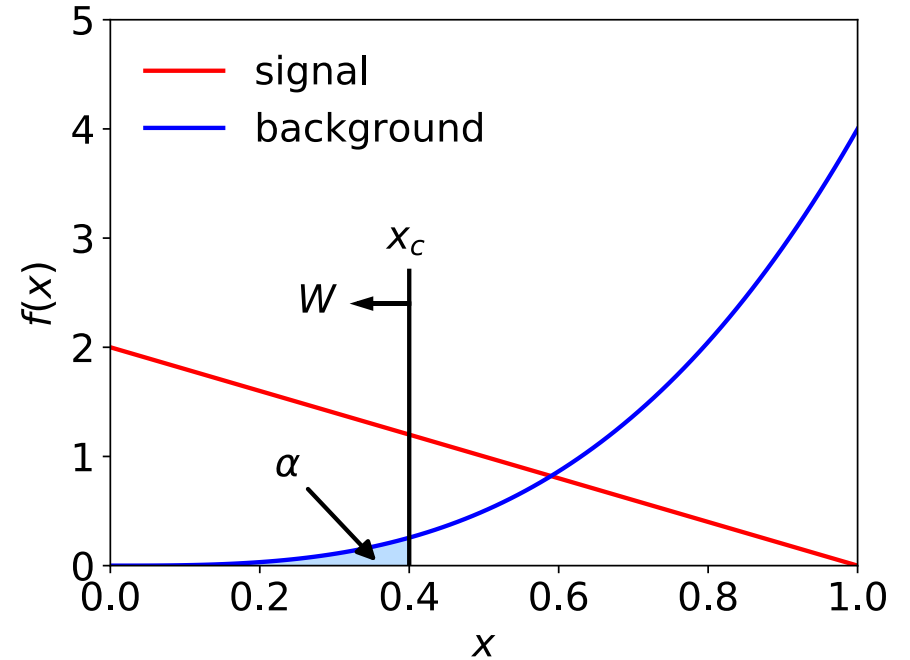
# Sınıflandırma için bir test örneği

Her olay için  $x$  niceliğini ölçebileceğimizi varsayalım, burada

$$f(x|s) = 2(1 - x)$$

$$f(x|b) = 4x^3$$

aralık  $0 \leq x \leq 1$ .



Sinyal (s) ve arka plan (b) karışımındaki her olay için test

$H_0$ : olay b türündedir

$W = \{ x : x \leq x_c \}$  biçiminde kritik bir bölge  $W$  kullanılarak, burada  $x_c$ , istenen  $\alpha$  büyüklüğünde bir test vermek için seçtiğimiz bir sabittir.

## Sınıflandırma örneği (2)

Diyelim ki  $\alpha = 10^{-4}$  istiyoruz. İfade edelim:

$$\alpha = P(x \leq x_c | b) = \int_0^{x_c} f(x|b) dx = \frac{4x^4}{4} \Big|_0^{x_c} = x_c^4$$

ve böylece  $x_c = \alpha^{1/4} = 0.1$

Bu test için (yani bu kritik bölge  $W$ ), sinyal hipotezine ( $s$ ) göre güç şudur:

$$P(x \leq x_c | s) = \int_0^{x_c} f(x|s) dx = 2x_c - x_c^2 = 0.19$$

Not: Optimal büyüklük ve güç, daha sonraki analizin hedeflerine bağlı olacak ayrı bir sorudur.

## Sınıflandırma örneği (3)

Bir olayın s veya b tipinde olma olasılığının ön olasılıkları şöyle olsun:

$$\pi_s = 0.001$$

$$\pi_b = 0.999$$

Seçilen sinyal örneğinin “saflığı” (b hipotezinin reddedildiği olaylar) Bayes teoremi kullanılarak bulunur:

$$\begin{aligned} P(s|x \leq x_c) &= \frac{P(x \leq x_c|s)\pi_s}{P(x \leq x_c|s)\pi_s + P(x \leq x_c|b)\pi_b} \\ &= 0.655 \end{aligned}$$

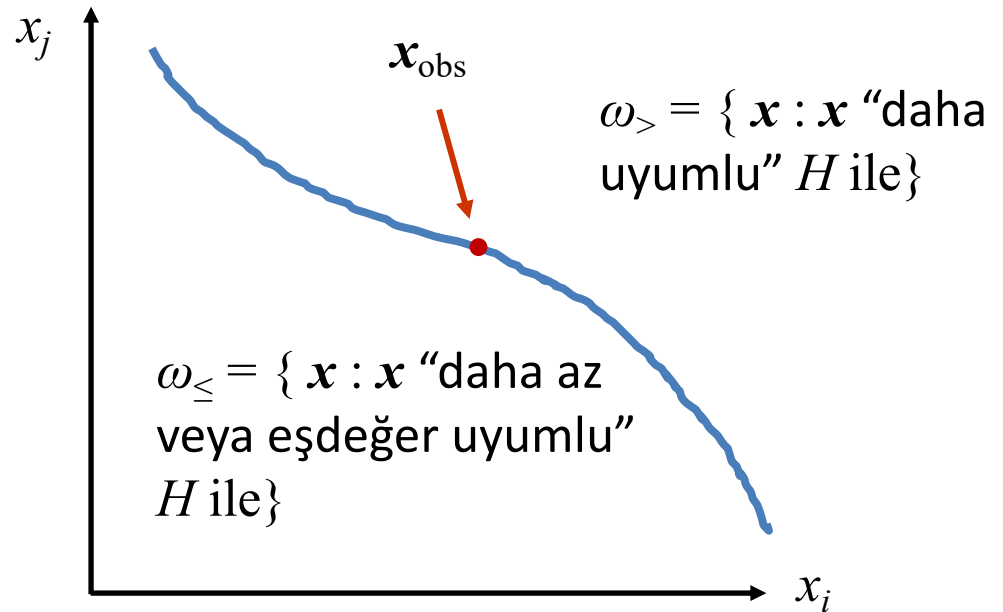
# Önem/uyum iyiliğinin test edilmesi

Varsayalım ki hipotez  $H$ ,  $x = (x_1, \dots, x_n)$  gözlem kümesi için pdf  $f(x|H)$ 'yi öngörüyor.

Bu uzayda tek bir nokta gözlemliyoruz:  $x_{\text{obs}}$ .

Veriler ile  $H$  öngörülleri arasındaki uyumluluk düzeyini nasıl ölçebiliriz?

Veri alanının hangi kısmının  $H$  ile  $x_{\text{obs}}$  noktasından daha eşit veya daha az uyumluluğu temsil ettiğine karar verin. (tek değil!)



# *p*-değerleri

Veriler ile hipotez arasındaki uyumluluk düzeyini (bazen 'uyum iyiliği')  $H$  için  $p$  değerini vererek ifade edin:

$$p = P(\mathbf{x} \in \omega_{\leq}(\mathbf{x}_{\text{obs}}) | H)$$

- =  $H$  varsayımı altında, elde ettiğimiz verilere göre  $H$  ile eşit veya daha az uyumlu verileri gözleme olasılığı.
- =  $H$  varsayımı altında, elde ettiğimiz verilerle  $H$  arasında en az aynı veya daha fazla uyumsuzluk gözleme olasılığı.

Temel fikir: daha iyi olmayan (veya eşit olan) uyumluluğa sahip veri bulma olasılığı çok düşükse,  $H$  "veriler tarafından tercih edilmez".

Eğer  $p$  değeri kullanıcı tarafından tanımlanan bir eşik  $\alpha$ 'nın (örneğin 0,05) altındaysa,  $H$  reddedilir (daha önce görüldüğü gibi  $\alpha$  büyüklüğündeki hipotez testine eşdeğerdir).

# *H*'nin *p*-değeri $P(H)$ değildir

H'nin *p* değeri H'nin doğru olma olasılığı değildir!

Frekansçı istatistiklerde  $P(H)$  hakkında konuşmayız (H tekrarlanabilir bir gözlemi temsil etmediği sürece).

Eğer  $P(H)$ 'yi tanımlarsak, örneğin Bayes istatistiklerinde bir inanç derecesi olarak, o zaman elde etmek için Bayes teoremini kullanmamız gerekir

$$P(H|\vec{x}) = \frac{P(\vec{x}|H)\pi(H)}{\int P(\vec{x}|H)\pi(H) dH}$$

burada  $\pi(H)$ , H için ön olasılıktır.

Şimdilik frekansçı yaklaşıma bağlı kalın;  
sonuç *p*-değeridir, ne yazık ki  $P(H)$  olarak yanlış da yorumlanabilir.

# Poisson sayma deneyi

Bir sayma deneyi yaptığımızı ve  $n$  tane olay gözlediğimizi varsayalım.

- \* olaylar sinyal sürecinden veya arka plandan kaynaklanabilir
- \* biz sadece toplam sayıyı sayıyoruz.

Poisson model:

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

$s$  = ortalama (yani, beklenen) # sinyal olayları

$b$  = ortalama # ardaalan olayları

Amaç  $s$  hakkında çıkarım yapmaktır, örneğin:

- \* test  $s = 0$  (red  $H_0 \approx$  “sinyal sürecinin keşfi”)
- \* test bütün sıfır-olmayan  $s$  (değerler red değil = güvenilirlik aralığı)

Her iki durumda ilgili alternatif hipotezin ne olduğunu sormak gerekir.



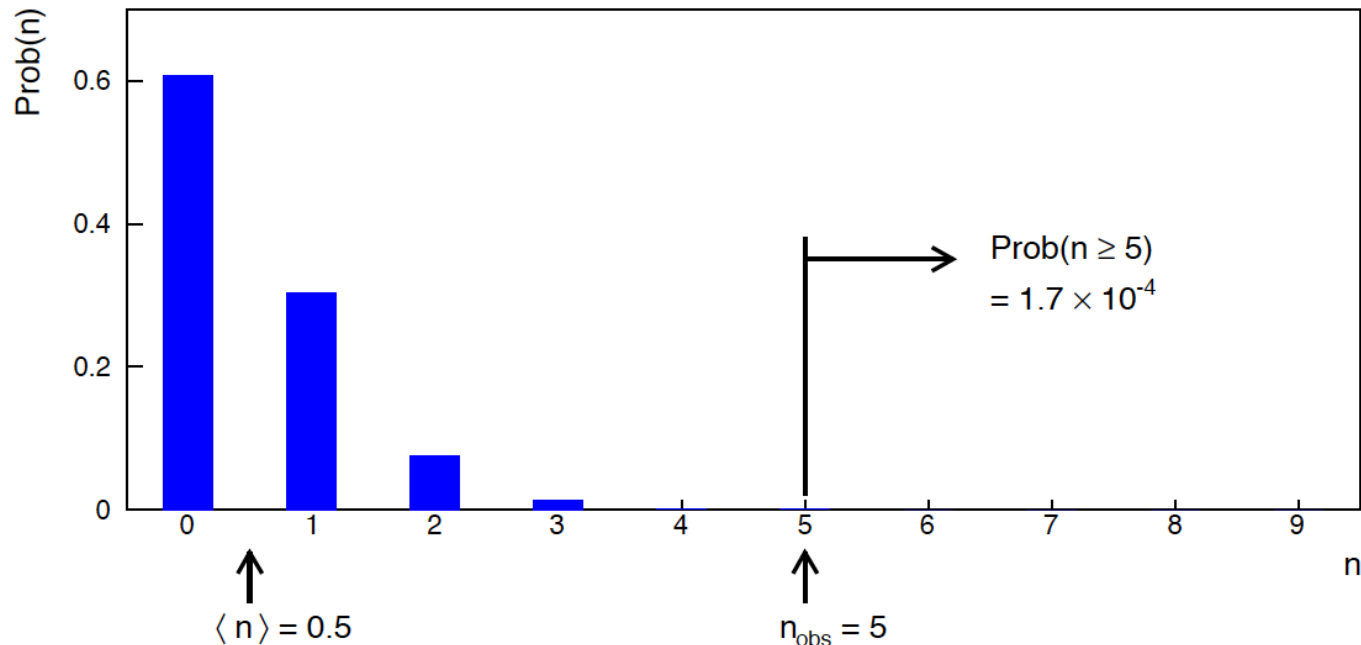
# Poisson sayma deneyi: keşif $p$ -değeri (value)

Varsayalım  $b = 0.5$  (bilinen), ve gözlem  $n_{\text{obs}} = 5$ .

Yeni bir keşif için kanıt talep etmeli miyiz?

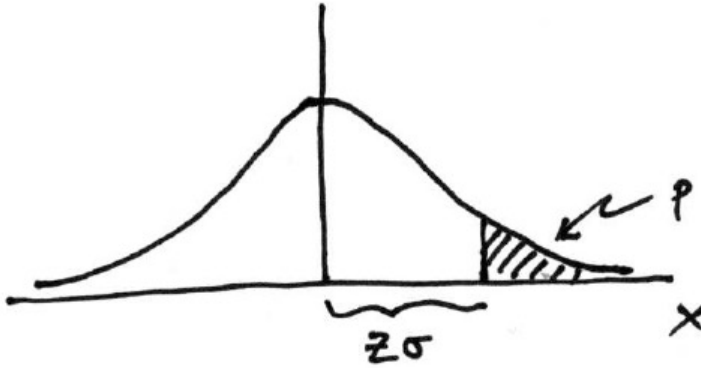
Hipotez  $s = 0$  için  $p$  değerini verin:

$$\begin{aligned} p\text{-value} &= P(n \geq 5; b = 0.5, s = 0) \\ &= 1.7 \times 10^{-4} \neq P(s = 0)! \end{aligned}$$



# *p*-değerinden önem

Genellikle  $Z$  önemini, bir Gauss değişkeninin aynı  $p$  değerini vermek için bir yönde dalgalanacağı standart sapma sayısı olarak tanımlarız.



$$p = \int_Z^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 1 - \Phi(Z)$$

$$Z = \Phi^{-1}(1 - p)$$

in ROOT:

```
p = 1 - TMath::Freq(Z)
```

```
Z = TMath::NormQuantile(1-p)
```

in python (scipy.stats):

```
p = 1 - norm.cdf(Z) = norm.sf(Z)
```

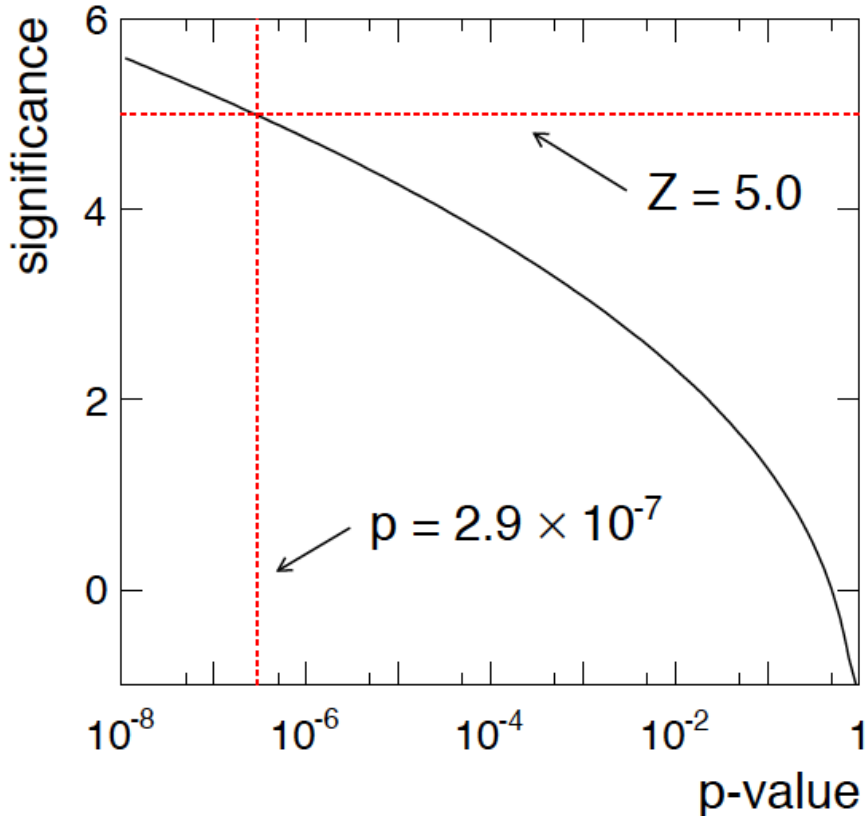
```
Z = norm.ppf(1-p)
```

Sonuç  $Z$  bir "sigma sayısı"dır. Bunun orijinal verinin Gauss dağılımına sahip olduğu anlamına gelmediğini unutmayın!

# Poisson sayma deneyi: keşif önemi

Burada  $p = 1.7 \times 10^{-4}$  için eşdeğer önem  $Z = \Phi^{-1}(1 - p) = 3.6$

$Z > 5$  ise genellikle keşif iddia edilir ( $p < 2.9 \times 10^{-7}$ , bir “5-sigma etkisi”)



Aslında bu geleneğin yeniden gözden geçirilmesi gerekir: p-değeri, yalnızca arka planı varsayarak sinyal benzeri bir dalgalanmanın olasılığını ölçmek için tasarlanmıştır; örneğin, gizli sistematik, makul sinyal modeli, verilerin sinyalle uyumluluğu, "başka yere bakma etkisi" (~çoklu test) vb. gibi durumları kapsamak için tasarlanmamıştır.

# Bir testi tersine çevirerek güven aralığı bulma

Bir parametrenin 'nokta tahminine' ek olarak, istatistiksel belirsizliğini yansıtan bir aralık bildirmeliyiz.

Bir parametre  $\theta$  için **güven aralıkları**, varsayılan değer  $\theta$ 'nin bir testini tanımlayarak bulunabilir (bunu tüm  $\theta$  için yapın):

$\theta$  tarafından 'dezavantajlı' olan verilerin değerlerini belirtin

(kritik bölge) böylece  $P(\text{kritik bölgedeki veriler} | \theta) \leq \alpha$

önceden belirlenmiş bir  $\alpha$  için, örneğin 0,05 veya 0,1.

Veriler kritik bölgede gözlemlenirse,  $\theta$  değerini reddedin.

Şimdi testi tersine çevirerek bir güven aralığı tanımlayalım:

$\alpha$  büyüklüğündeki bir testte reddedilmeyen  $\theta$  değerleri kümesi

(güven düzeyi (CL)  $1 - \alpha$ 'dır).

# Güven aralığı ile p değeri arasındaki ilişki

Eşdeğer olarak, her varsayılan  $\theta$  değeri için bir anlamlılık testi düşünebiliriz ve bu da bir p-değeri olan  $p_\theta$  ile sonuçlanır.

eğer  $p_\theta \leq \alpha$ , o zaman  $\theta$  red edilir.

CL =  $1 - \alpha$ 'daki güven aralığı, reddedilmeyen  $\theta$  değerlerinden oluşur. Örneğin,  $\theta$ 'deki bir üst sınır,  $p_\theta > \alpha$  olan en büyük değerdir.

Pratikte  $p_\theta = \alpha$  koyarak bulun ve  $\theta$  için çözün.

Çok boyutlu bir parametre uzayı  $\theta = (\theta_1, \dots, \theta_M)$  için aynı fikri kullanın – sonuç, sınırı  $p_\theta = \alpha$  tarafından belirlenen bir güven "bölgesi"dir.

# Güven aralığının kapsam olasılığı

$\theta$ 'nin gerçek değeri reddedilirse, o zaman güven aralığında değildir. Bunun olasılığı, yapılandırmaya göredir (sürekli veriler için eşitlik):

$$P(\text{red } \theta | \theta) \leq \alpha = \text{tip-I hata oranı}$$

Bu nedenle, aralığın  $\theta$ 'yi içermesi veya "örtmesi" olasılığı şudur:

$$P(\text{güvenilirlik aralığı "kapsar" } \theta | \theta) \geq 1 - \alpha$$

Bu, dikkate alınan  $\theta$  değerleri kümesinin gerçek değeri içerdiğini, yani bileşik hipotez  $P(x | H, \theta)$  olduğunu varsayar.

# Poisson parametresi üzerine frekansçı üst limit

$n \sim \text{Poisson}(s + b)$  gözlemlene durumu tekrar düşünün.

Varsayalım ki  $b = 4,5$ ,  $n_{\text{obs}} = 5$ .

Burada  $s$ 'nin %95 CL'deki üst sınırını bulun.

İlgili alternatif  $s = 0$ 'dır (düşük  $n$ 'deki kritik bölge)

Varsayımsal  $s$ 'nin  $p$  değeri  $P(n \leq n_{\text{obs}}; s, b)$ 'dir

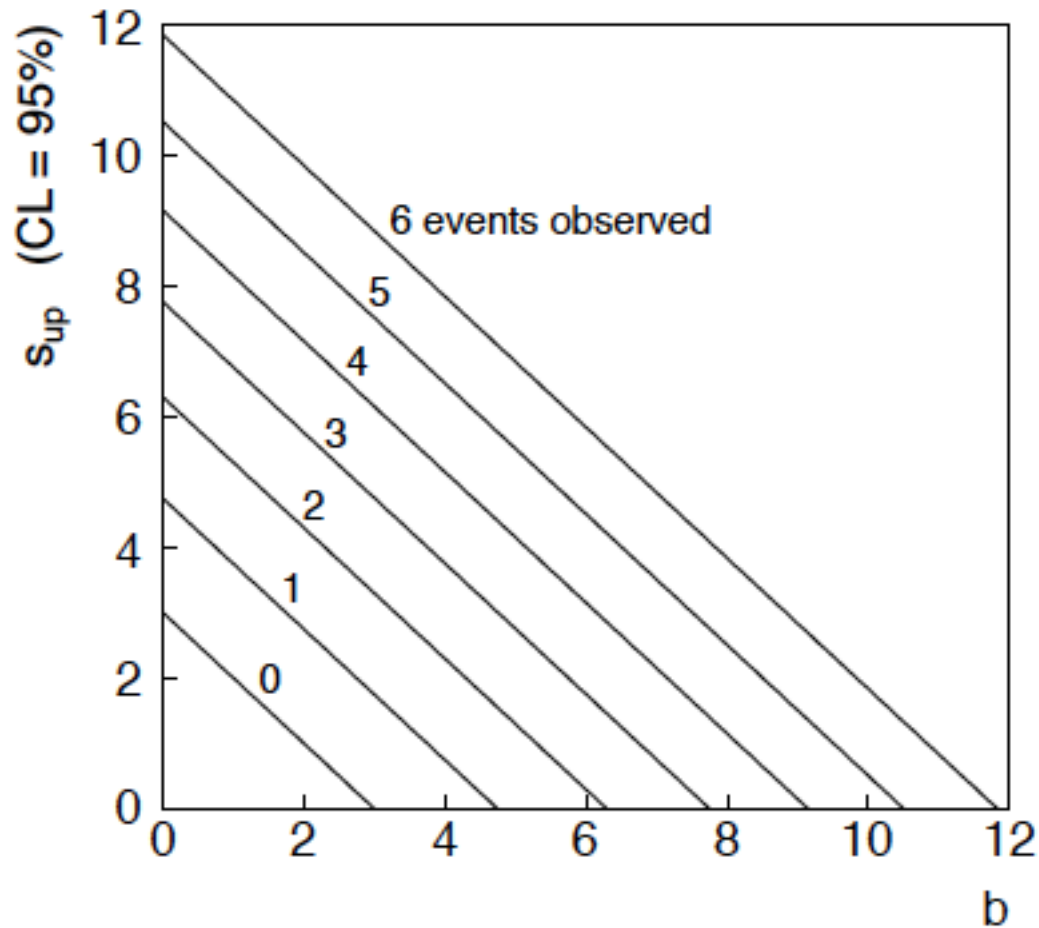
CL = 1'deki üst sınır –  $\alpha$ 'dan bulunur

$$\alpha = P(n \leq n_{\text{obs}}; s_{\text{up}}, b) = \sum_{n=0}^{n_{\text{obs}}} \frac{(s_{\text{up}} + b)^n}{n!} e^{-(s_{\text{up}} + b)}$$

$$\begin{aligned} s_{\text{up}} &= \frac{1}{2} F_{\chi^2}^{-1}(1 - \alpha; 2(n_{\text{obs}} + 1)) - b \\ &= \frac{1}{2} F_{\chi^2}^{-1}(0.95; 2(5 + 1)) - 4.5 = 6.0 \end{aligned}$$

# $n \sim \text{Poisson}(s+b)$ : $s$ üzerine frekansçı üst limit

$n$ 'nin düşük dalgalanması için formül,  $s_{\text{up}}$  için negatif sonuç verebilir; yani güven aralığı boştur;  $s \geq 0$ 'ın tüm değerleri  $p_s \leq \alpha$ 'dır.





# Parametre uzayının sınırına yakın limitler

Varsayım, örneğin,  $b = 2.5$  ve gözlem  $n = 0$ .

Eğer  $CL = 0.9$  seçersek,  $s_{up}$  için aşağıdaki sonucu buluruz

$$s_{up} = -0.197 \quad (CL = 0.90)$$

## Fizikçi:

Başlamadan önce  $s \geq 0$  olduğunu biliyorduk; pahalı deneyin sonucunu bildirmek için negatif üst sınır kullanamayız!

## İstatistikçi:

Aralık, gerçek değeri yalnızca %90 oranında kapsayacak şekilde tasarlanmıştır — bu açıkça böyle zamanlardan biri değildi.

Çok az deneysel duyarlılığa sahip olunan parametre değerlerinin (örneğin çok küçük  $s$ ) test edilmesi sırasında karşılaşılan nadir bir ikilem değildir.

# s=0 için beklenen limit

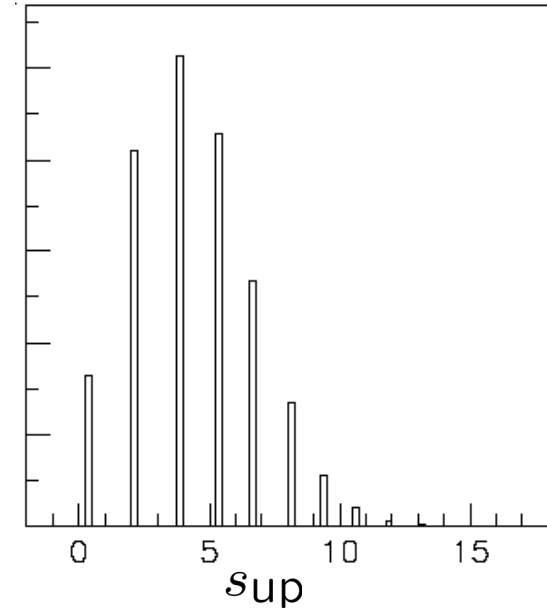
Fizikçi yorumu: CL = 0.95 kullanmalıyım — sonra  $s_{up} = 0.496$

Daha iyisi: CL = 0.917923 için  $s_{up} = 10^{-4}$  alırız !

Gerçeklik kontrolü:  $b = 2,5$  ile,  $n$ 'deki tipik Poisson dalgalanması en azından  $\sqrt{2,5} = 1,6$ 'dır. Sınır nasıl bu kadar düşük olabilir?

Sinyal yok hipotezi için  
ortalama limite bakın ( $s = 0$ )  
(duyarlılık).

$b = 2,5$  ,  $s = 0$  ile  
%95 CL sınırlarının dağılımı.  
Ortalama üst sınır = 4,44



# Ekler

# Olasılık fonksiyonundan yaklaşık güven aralıkları/bölgeleri

Parametre değerini (değerlerini)  $\theta = (\theta_1, \dots, \theta_n)$  oranını kullanarak test ettiğimizi varsayalım

$$\lambda(\theta) = \frac{L(\theta)}{L(\hat{\theta})} \quad 0 \leq \lambda(\theta) \leq 1$$

Daha düşük  $\lambda(\theta)$ , veriler ile varsayılan  $\theta$  arasındaki daha kötü uyum anlamına gelir. Eşdeğer olarak, genellikle şunu tanımlar:

$$t_\theta = -2 \ln \lambda(\theta)$$

Daha yüksek  $t_\theta$ ,  $\theta$  ile veriler arasındaki uyumun daha kötü olduğu anlamına gelir.

$\theta$ 'nin p değeri bu nedenle,  $p_\theta = \int_{t_{\theta, \text{obs}}}^{\infty} f(t_\theta | \theta) dt_\theta$

pdf ihtiyacı

# Güven bölgesi

Wilk teoremi (büyük örneklem limitinde ve belirli koşullar sağlandığında...) ifade eder.

$$f(t_{\theta}|\theta) \sim \chi_n^2$$

ki-kare dağılımı #d.o.f. =  
#bileşenler  $\theta = (\theta_1, \dots, \theta_n)$ .

Bunun geçerli olduğunu varsayarsak, p değeri şu şekildedir:

$$p_{\theta} = 1 - F_{\chi_n^2}(t_{\theta}) \quad \leftarrow \alpha \text{ ya eşitlenir}$$

Güven bölgesinin sınırını bulmak için  $p_{\theta} = \alpha$ 'yı ayarlayın ve  $t_{\theta}$ 'yi çözün:

$$t_{\theta} = F_{\chi_n^2}^{-1}(1 - \alpha)$$

ve bunu hatırlayın  $t_{\theta} = -2 \ln \frac{L(\theta)}{L(\hat{\theta})}$

## Güven bölgesi (devam)

Burada  $\theta$  uzayındaki güven bölgesinin sınırı şudur:

$$\ln L(\theta) = \ln L(\hat{\theta}) - \frac{1}{2} F_{\chi_n^2}^{-1}(1 - \alpha)$$

Örneğin,  $1 - \alpha = \%68,3$  ve  $n = 1$  parametresi için,

$$F_{\chi_1^2}^{-1}(0.683) = 1$$

ve böylece  $\%68,3$  güven aralığı şu şekilde belirlenir:

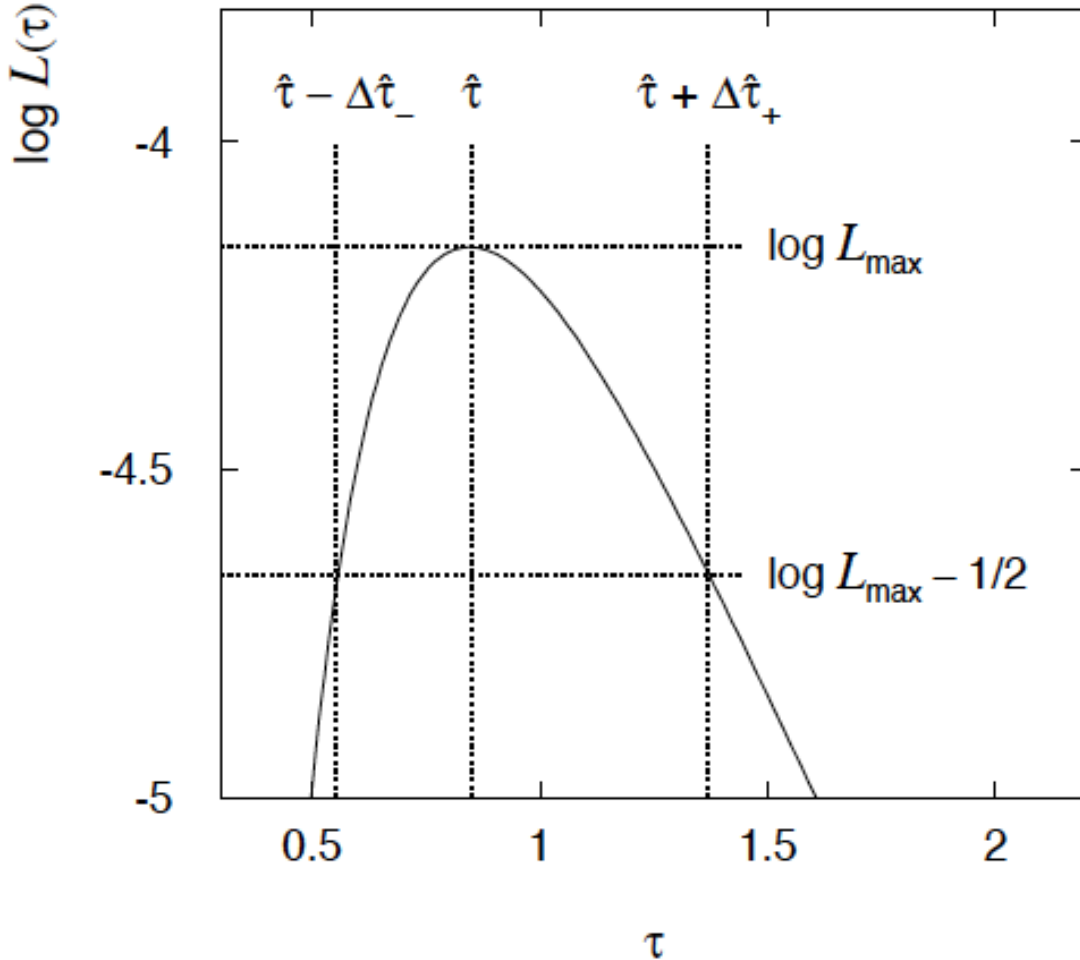
$$\ln L(\theta) = \ln L(\hat{\theta}) - \frac{1}{2}$$

Tahmincinin standart sapmasını bulma tarifiyle aynıdır, yani,

$[\hat{\theta} - \sigma_{\hat{\theta}}, \hat{\theta} + \sigma_{\hat{\theta}}]$   $\%68,3$  CL güven aralığıdır.

# In L( $\theta$ ) aralığından bir örnek

n=1 parametresi için CL = 0,683,  $Q_\alpha = 1$ .



Üstel örneğimiz, artık yalnızca  $n = 5$  olay ile.

ML tahminini

In  $L_{\max} - 1/2$ 'den yaklaşık güven aralığıyla "asimetrik hata çubuğu" olarak raporlayabiliriz:

$$\hat{\tau} = 0.85_{-0.30}^{+0.52}$$

# Çok parametrelili durum

Parametre sayısının artmasıyla,  $CL = 1 - \alpha$ , verilen bir güven bölgesi tarafından belirlenen değer için azalır.

$$Q_\alpha = F_{\chi_n^2}^{-1}(1 - \alpha)$$

$Q_\alpha$	$1 - \alpha$				
	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
1.0	0.683	0.393	0.199	0.090	0.037
2.0	0.843	0.632	0.428	0.264	0.151
4.0	0.954	0.865	0.739	0.594	0.451
9.0	0.997	0.989	0.971	0.939	0.891



# Çok parametrelili durum (devam)

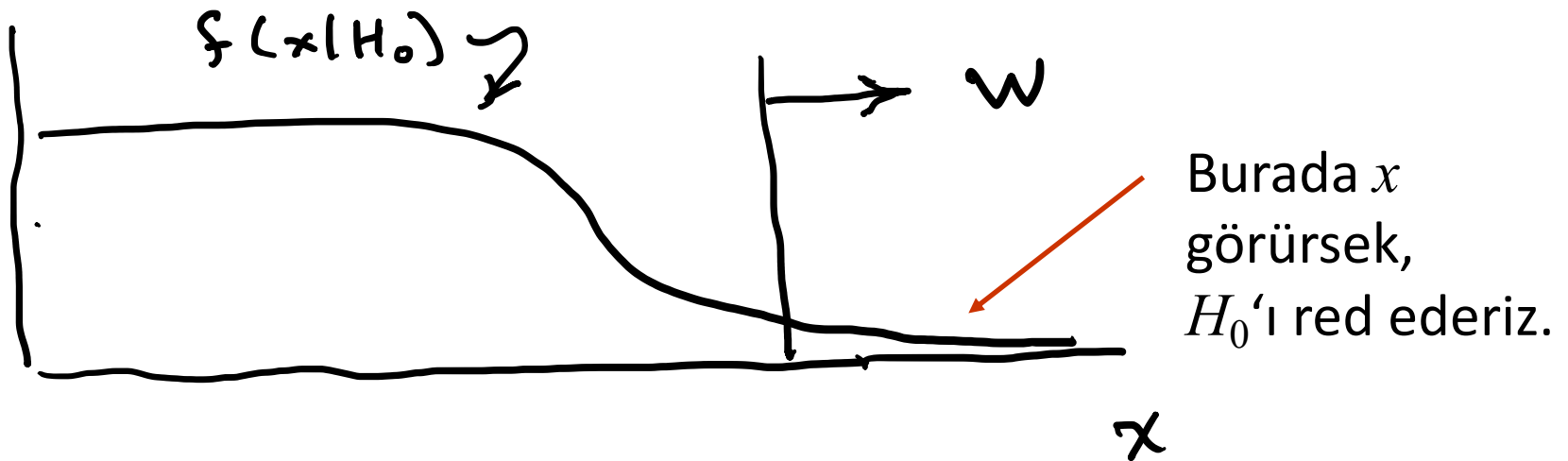
Eşdeğer olarak,  $Q_\alpha$ , verilen bir  $CL = 1 - \alpha$  için  $n$  ile artar.

$1 - \alpha$	$\tilde{Q}_\alpha$				
	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
0.683	1.00	2.30	3.53	4.72	5.89
0.90	2.71	4.61	6.25	7.78	9.24
0.95	3.84	5.99	7.82	9.49	11.1
0.99	6.63	9.21	11.3	13.3	15.1

# W'yi nereye uygulayacađımız belli mi?

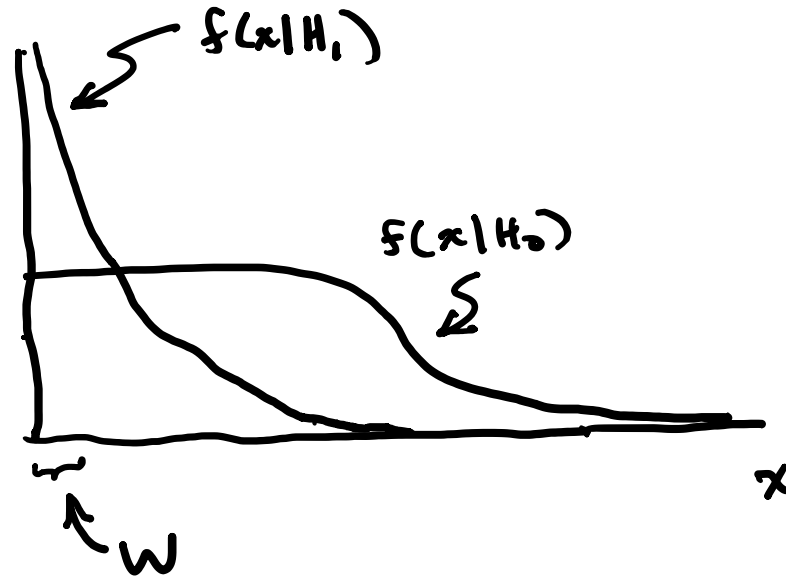
1930'larda alternatif hipotezin rolü konusunda büyük tartiřmalar vardı. Fisher, bir alternatifte atıfta bulunmadan bir hipotez  $H_0$ 'ın test edilebileceđini savundu.

Örneđin,  $H_0$ 'ın  $x$ 'in (pozitif olduđunu varsayalım) genellikle düşük çıkacađını öngördüđünü varsayalım. Yüksek  $x$  deđerleri  $H_0$  için daha az karakteristiktir, bu nedenle yüksek bir deđer gözlemlenirse  $H_0$ 'ı reddetmeliyiz, yani W'yi yüksek  $x$ 'e uygulamalıyız:



# W'yi nereye uygulayacağımız belirgin değil mi?

Durum:  $H_0$ 'a ilişkin tek geçerli alternatif aşağıdaki gibi  $H_1$  ise



Burada yüksek  $x$ ,  $H_0$ 'ın daha karakteristik özelliğidir ve  $H_1$ 'den beklediğimiz gibi değildir. Bu yüzden  $W$ 'yi düşük  $x$ 'e koymak daha iyidir. Neyman ve Pearson, " $H_0$ 'ın daha az karakteristik özelliği"nin yalnızca "ilgili bazı alternatif  $H_1$ 'in daha fazla karakteristik özelliği" anlamına geldiği zaman iyi tanımlandığını savunmuştur.

# Tip-I, Tip-II hatalar

$H_0$  hipotezini doğru olduğunda reddetmek Tip-I hatasıdır.

Bunun için maksimum olasılık testin boyutudur:

$$P(x \in W | H_0) \leq \alpha$$

Ancak  $H_0$  yanlış olduğunda ve alternatif  $H_1$  doğru olduğunda da kabul edebiliriz. Buna Tip II hatası denir ve olasılıkla gerçekleşir

$$P(x \in S - W | H_1) = \beta$$

Bir eksi bu ifade, testin alternatif  $H_1$ 'e göre gücü olarak adlandırılır:

$$\text{Güç (power)} = 1 - \beta$$

# p-değerinin dağılımı

p-değeri, verilerin bir fonksiyonudur ve dolayısıyla kendisi de belirli bir dağılıma sahip rastgele bir değişkendir.

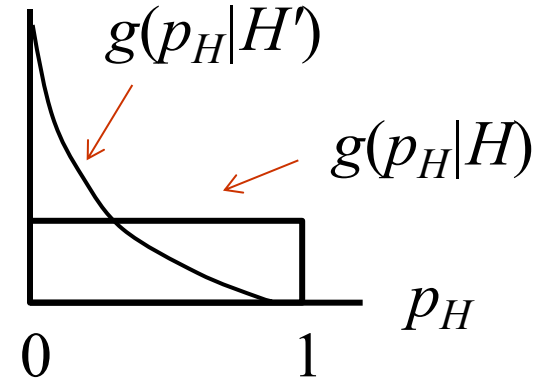
H'nin p-değerinin bir test istatistiği  $t(x)$ 'ten şu şekilde bulunduğunu varsayalım:

$$p_H = \int_t^{\infty} f(t'|H) dt'$$

H varsayımı altında  $p_H$ 'in pdf'si

$$g(p_H|H) = \frac{f(t|H)}{|\partial p_H / \partial t|} = \frac{f(t|H)}{f(t|H)} = 1 \quad (0 \leq p_H \leq 1)$$

Genel olarak sürekli veriler için,  
H varsayımı altında  $p_H \sim \text{Uniform}[0,1]$   
'dir ve bazı (geniş) alternatif sınıfları  
için sıfıra doğru yoğunlaşır.

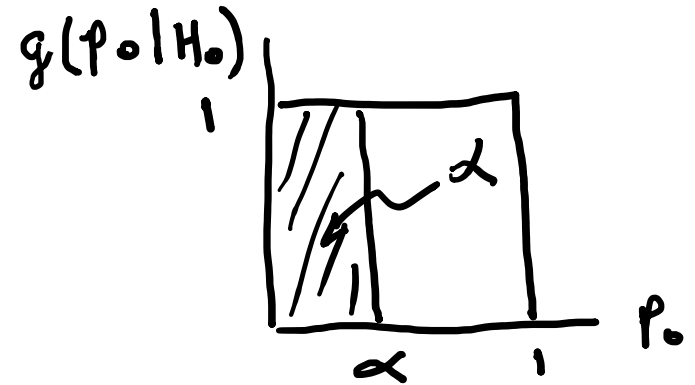


# H0 testini tanımlamak için bir p değeri kullanma

Bir hipotez  $H_0$  varsayımı altında, onun p-değeri  $p_0$ 'ın  $[0,1]$ 'de düzgün bir dağılım izlediği gösterilebilir.

Yani  $p_0$ 'ı verilen  $\alpha$ 'dan daha az bulma olasılığı şudur:

$$P(p_0 \leq \alpha | H_0) = \alpha$$



Yani  $\alpha$  büyüklüğündeki bir  $H_0$  testinin kritik bölgesini,  $p_0 \leq \alpha$  olan veri alanı kümesi olarak tanımlayabiliriz.

Formal p değeri yalnızca  $H_0$  ile ilgilidir, ancak ortaya çıkan test, belirli bir alternatif  $H_1$ 'e göre belirli bir güce (power) sahip olacaktır.



# Kaynaklar

- [1] S. Navas *et al.* (Particle Data Group), *Statistics* Revised by G. Cowan, Phys. Rev. D 110, 030001 (2024).
- [2] L. Lista, *Statistical Methods for Data Analysis in Particle Physics*, (Springer, 2017).
- [3] I. Narsky, F.C. Porter, *Statistical Analysis Techniques in Particle Physics*, (Wiley-Vch, Germany, 2014).
- [4] O. Behnke, K. Kroninger, G. Schott, and T. Schorner-Sadenius, *Data Analysis in High Energy Physics*, (Wiley-Vch, Germany, 2013).
- [5] F. James, *Statistical Methods in Experimental Physics*, 2nd ed. (World Scientific, Singapore, 2006).
- [6] S. Brandt, *Data Analysis*, 3rd ed. (Springer, New York, 1999).
- [7] G.D. Cowan, *Statistical Data Analysis*, (Oxford University Press, 1998).
- [8] R.J. Barlow, *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences* (Wiley, Chichester, 1989).
- [9] L. Lyons, *Statistics for Nuclear and Particle Physicists*, (Cambridge University Press, 1986).