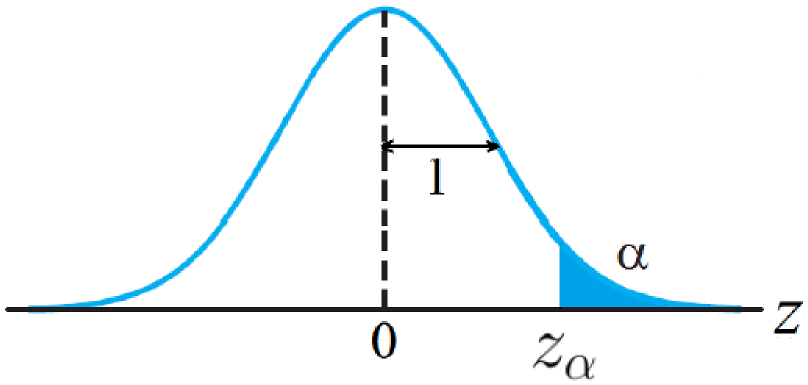
İstatistik Analiz (Keşif, Dışarlama ve Limit Koyma)



Ahmet Bingül
ODTÜ Fizik Bölümü
abingul@metu.edu.tr

İÇERİK

- **Giriş**
- **Örnekleme**
- **Hipotez Testleri**
- **Alıştırma**

Giriş

Parçacık Fiziği’nde "gerçek fiziksel model" hakkında çıkarımlarda bulunurken hipotez testiyle uğraşırız.

Deneysel veriler göz önüne alındığında bir karar vermek gerekir.
(örneğin dışarlama ya da keşif)

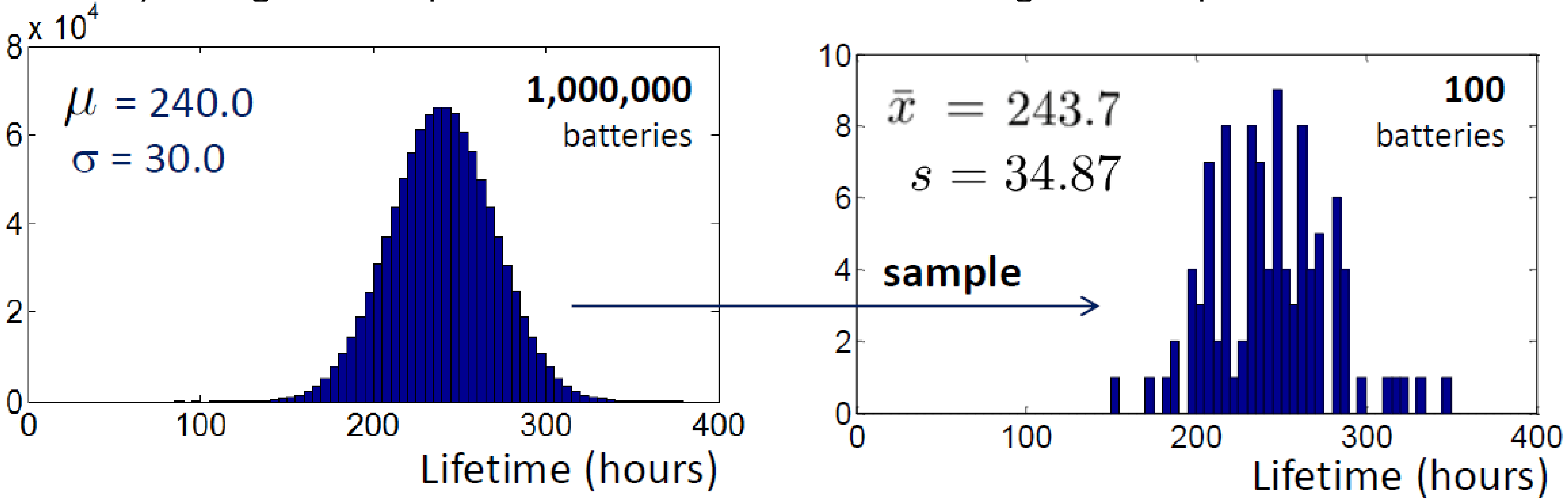
Rastgele Örnekleme

Anakütle (Popülasyon) = ilgilendiğimiz gözlemlerin hepsi

Örnek = anakütlenin altkümesi

Poülasyon Dağılımı: 1M pilin ömrü

Örnek Dağılımı: 100 pilin ömrü

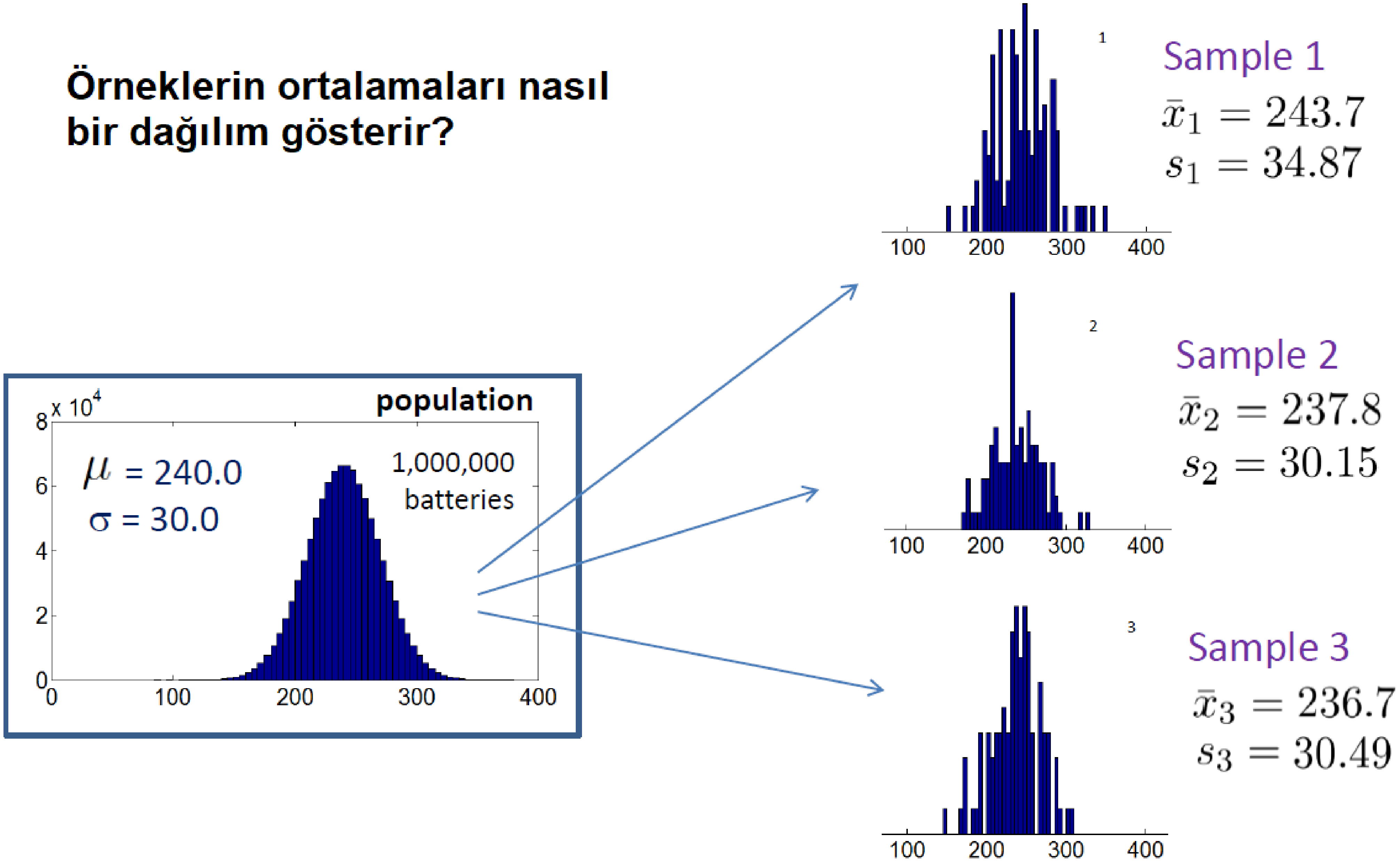


Anakütlenin ortalaması ve standart sapması, örnek'ten çıkarılabilir. Ancak, bu çıkarımda bazı belirsizlikler vardır.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

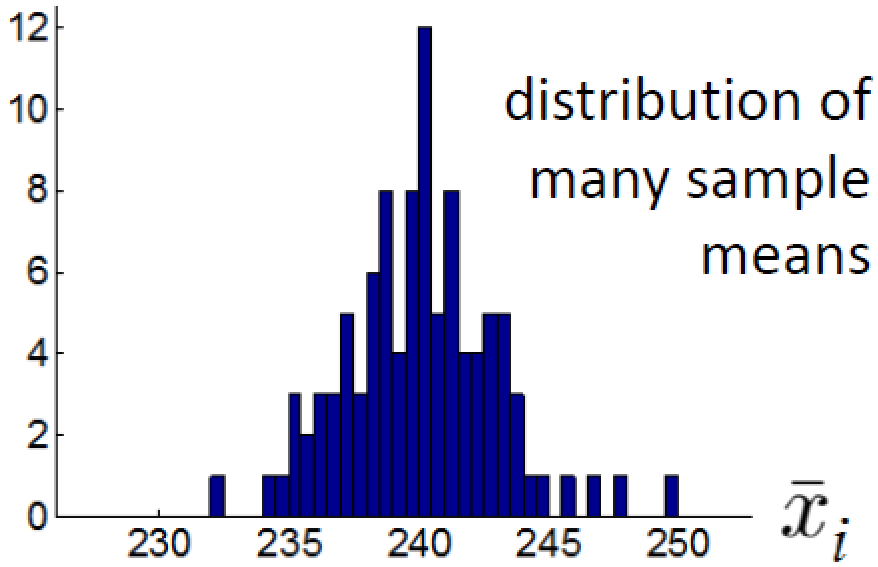
Rastgele Örnekleme

Örneklerin ortalamaları nasıl bir dağılım gösterir?



Rastgele Örnekleme

Her biri n = 100 gözlem içeren, N = 100 tane farklı örnek alalım. Her bir örneğin ortalaması $\bar{x}_i$ olsun.



Gözlem sayısı arttıkça ortalamaların ortalaması

$$\mu_{\bar{X}} \to \mu$$

standart sapması:

$$\sigma_{\bar{X}} \to \sigma/\sqrt{n}$$

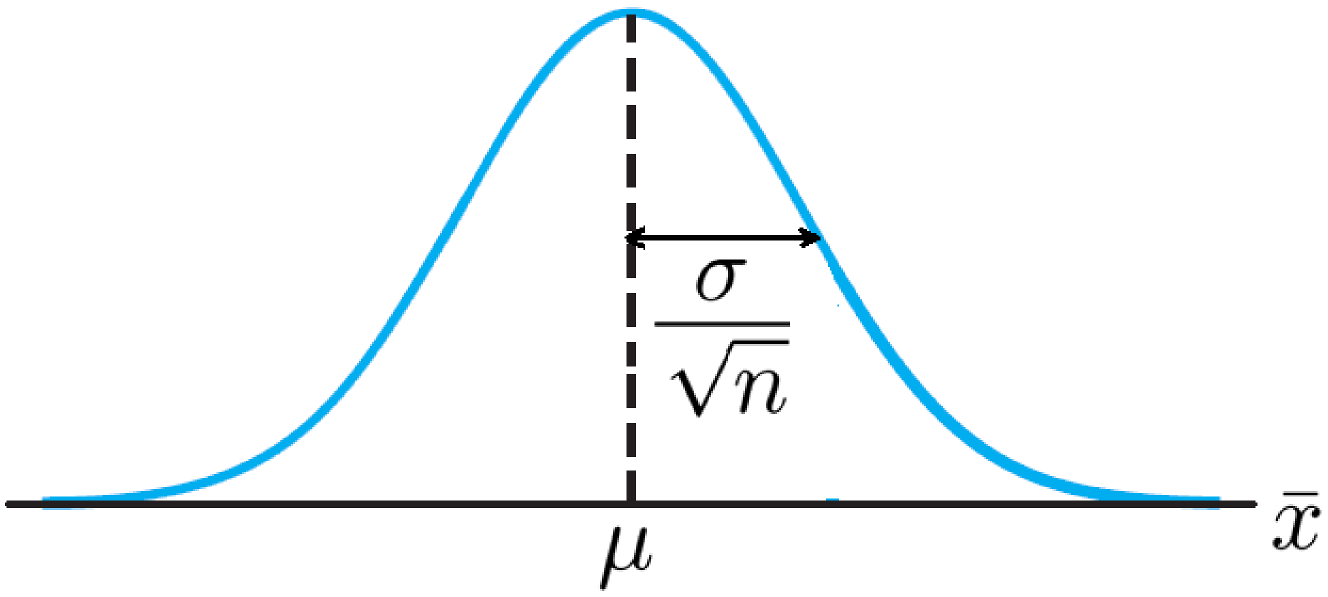
olur.

$$\mu_{\bar{X}} = \frac{1}{N}\sum_{i=1}^{N} \bar{x}_i = 240.05 \approx \mu$$

$$\sigma_{\bar{X}} = 2.9 \approx \frac{\sigma}{\sqrt{n}} = 3$$

Merkezi limit teoremi:

Her biri *n* tane gözlem içeren örneklerin ortalamalarının dağılımı, popülasyonun ortalaması değerine etrafında standart sapması $\sigma/\sqrt{n}$ olacak şekilde dağılır.

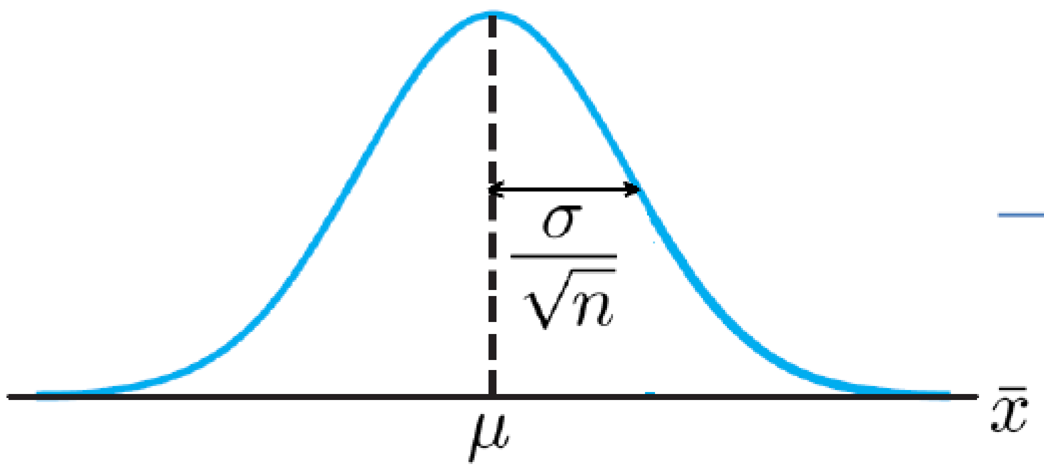


$$\mu_{\bar{X}} \to \mu \quad \text{and} \quad \sigma_{\bar{X}} \to \sigma/\sqrt{n}$$

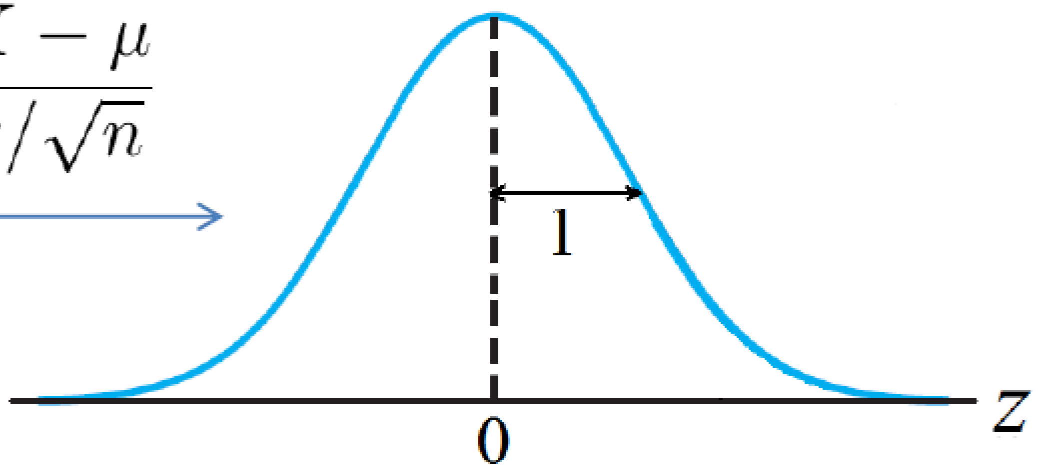
Bu ortalamanın **standart hatası** olarak bilinir.

Z-Dönüşümü

n yeterince büyükse, *Z* değeri standart normal olarak dağılır.



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

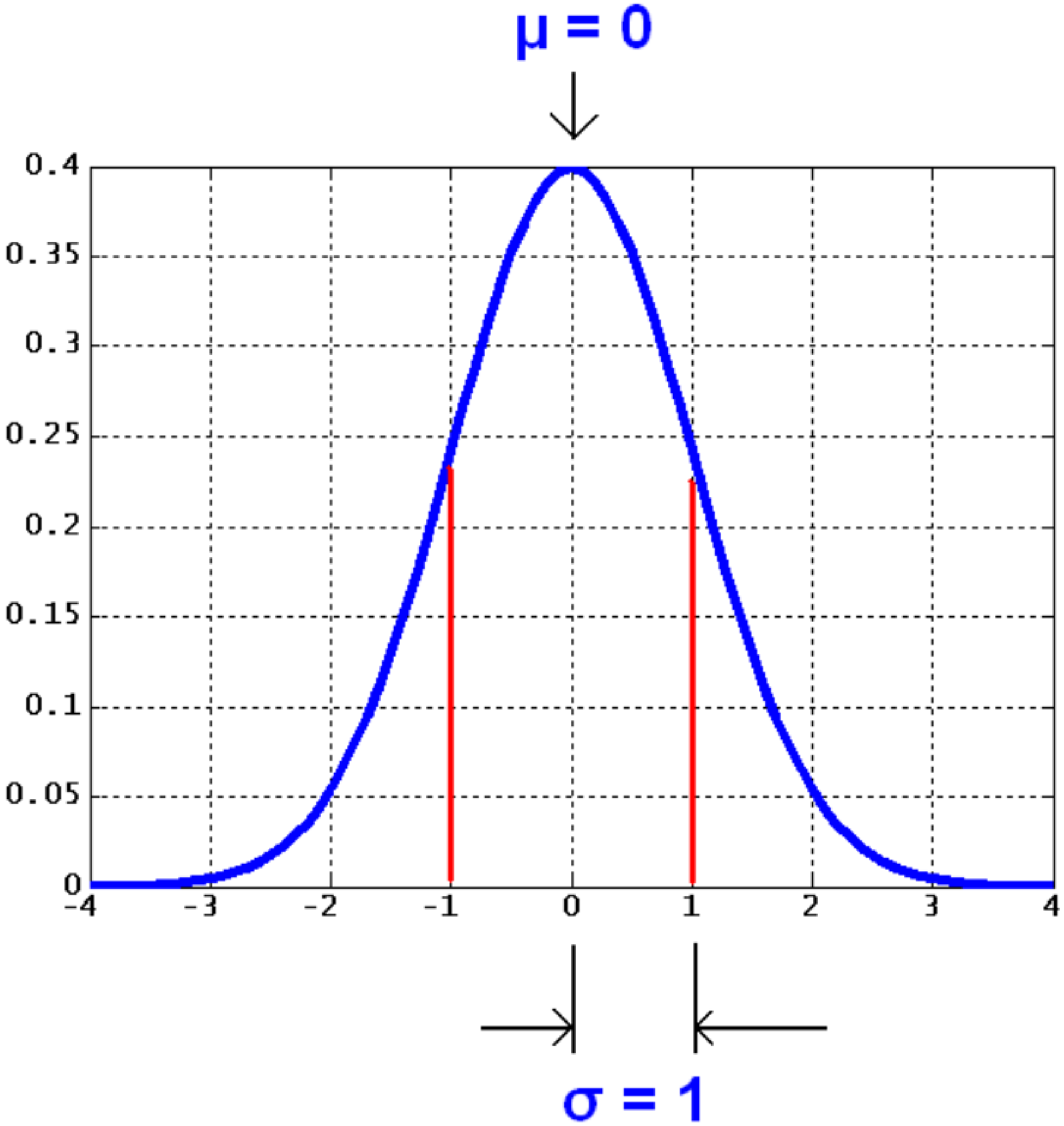


The distribution of $\bar{x}$ with large n.

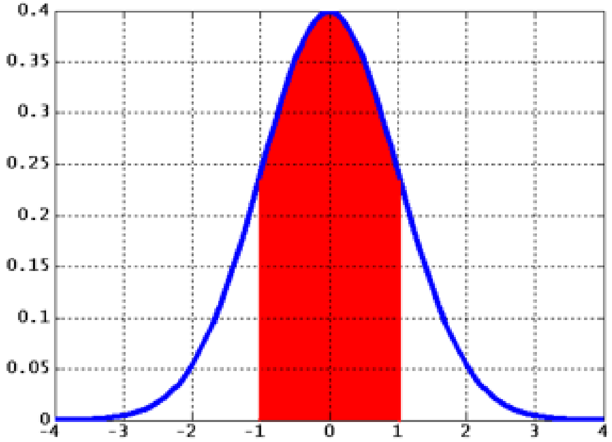
The Z transformed distribution.

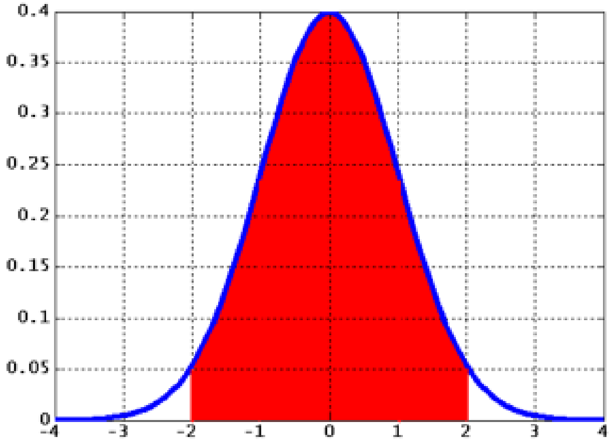
$$\int_a^b f(z)dz = \Phi(b) - \Phi(a) \longleftarrow f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

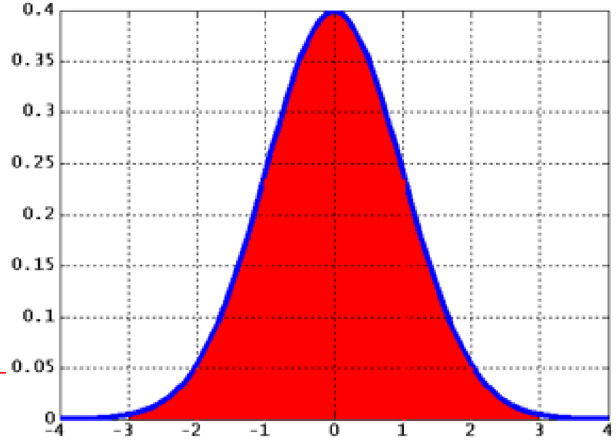
Standard Normal Curve

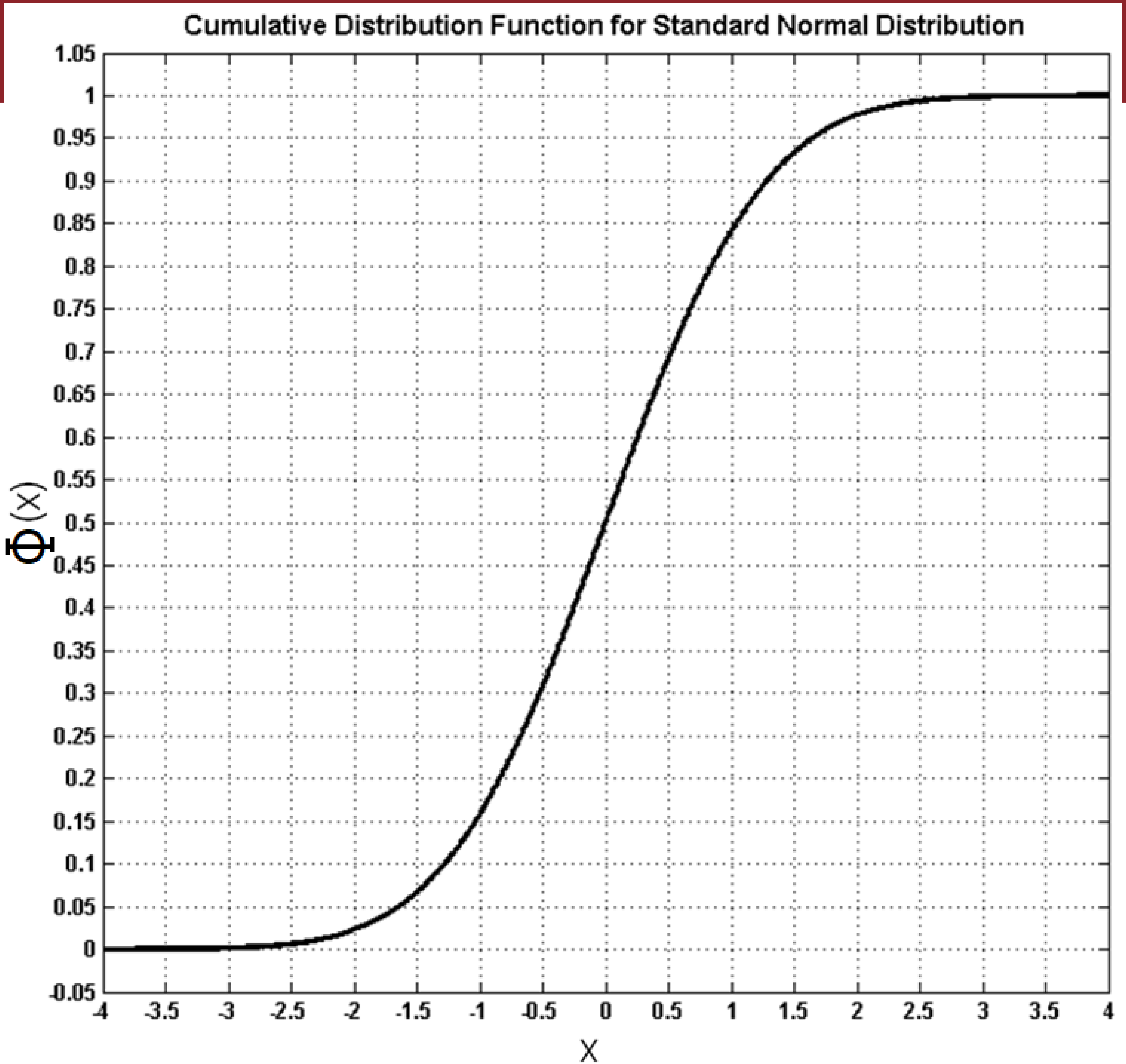


CL	Eğri altında kalan alan
%68	$\int_{-1}^{1} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 0.6827$
%95	$\int_{-2}^{2} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 0.9545$
%99.7	$\int_{-3}^{3} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 0.9973$



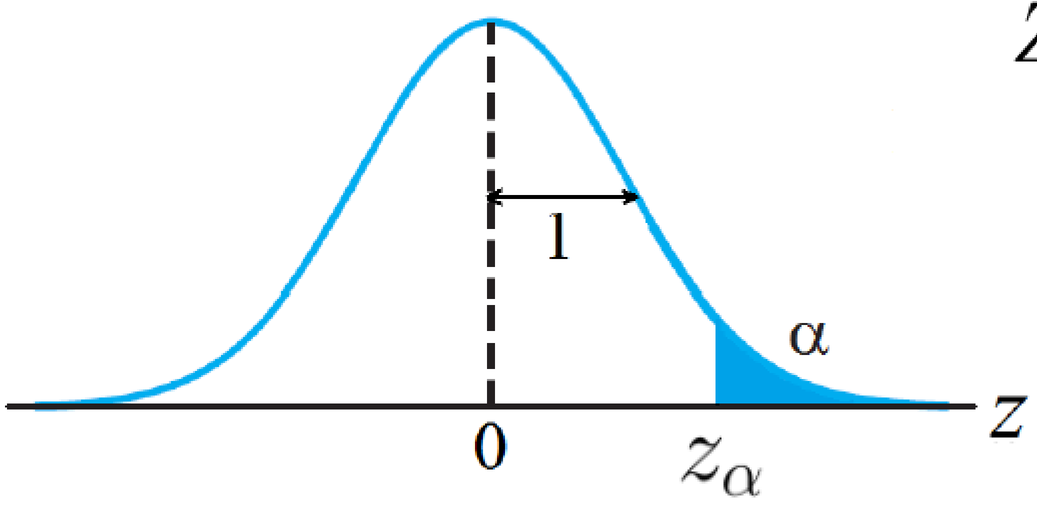





Cumulative Distribution Function for Standard Normal Distribution
1.05
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
Φ(x)
-4
-3.5
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
3
3.5
4
X

Örnek Olasılığı

$\alpha = z_\alpha$ 'dan sonraki alan:



$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \longrightarrow P\left(Z > z_\alpha\right) = \alpha$$

$$P(Z > z_\alpha) = \alpha = \int_{z_\alpha}^{\infty} f(z)dz = \Phi(\infty) - \Phi(z_\alpha)$$

$$= 1.0 - \Phi(z_\alpha)$$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

$\Phi(x) =$ **`ROOT::Math::normal_cdf(x)`**

Hipotez Testi (P-değeri Yaklaşımı)

Genel Fikir:

1. Bir varsayım yap (hipotez)
2. Varsayımlarla ilgili gözlem yap (veri topla)
3. P-değerine dayalı, alternatif hipotezin ne kadar güçlü bir şekilde önerildiğini belirle. P-değeri bir eşik değerden küçükse hipotez reddedilir (karar ver).

Mahkemede, sanığın başlangıçta suçsuz olduğu varsayılır:

Sıfır Hipotezi: H_0 = suçsuz

Alternatif Hipotez: H_A = suçlu

Eğer **P(suç kanıtına dair gözlemler | suçsuz) = makul derecede küçük değer**

ise => sanık suçlu kabul edilir (H_0 reddedilir).

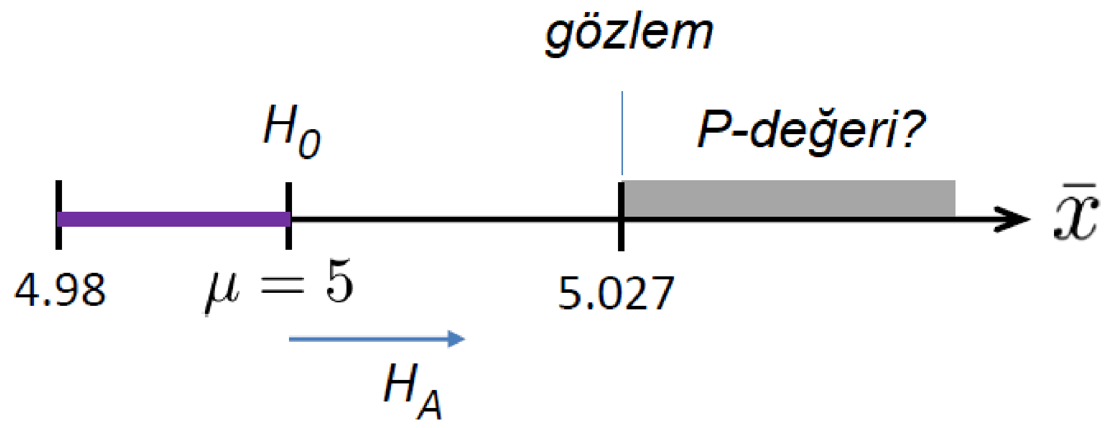
Örnek

Kararsız bir parçacığın kütlesi, detektör etkilerinden dolayı, ortalaması 4.99 ± 0.01 GeV standart sapması 0.12 GeV olacak şekilde normal dağılmaktadır. Bir öğrenci büyük bir kümeden 100 tane parçacık içeren bir örnek seçiyor ve örnek dağılımın ortalamasını 5.027 GeV olarak tespit ediyor. Parçacığın kütlesi 5 GeV olduğu varsayımı doğru mudur?

$$H_0: \mu = 5$$

$$H_A: \mu > 5$$

$$\bar{x} = 5.027$$



$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{5.027 - 5.00}{0.12/\sqrt{100}} = 2.25$$

$$\text{P-değeri} = \alpha = P(Z > 2.25)$$
$$= 1 - \Phi(2.25) = 0.012$$

P-değeri = %1.2
Bu bize güçlü bir şekilde $\mu > 5$ GeV olduğunu belirtiyor.
H_0 hipotezini reddediyoruz.
Ancak, %1.2 olasılıkla bu kararımızda yanılabiliriz.

Hipotez Testi (P-değeri Yaklaşımı)

Hipotez testleri, p-değeri yerine bazen Z değeri üzerinden değerlendirilebilir.

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

Z önem (significance) olarak adlandırılır.

1-p değeri Güvenlik Düzeyi (Confidence Level) olarak adlandırılır.

p	CL = 1-p	Z
0.10	0.90	1.282
0.05	0.95	1.645
0.01	0.99	2.326
2.87×10^{-7}	~1.00	5.000

- **Z = 1.645**, %95'lik güvenlik düzeyinde tek taraflı limit koymak için kullanılır.
- **3 < Z < 5** verinin yeni bir keşfi ortaya koyabileceğinin kanıtıdır.
- **Z > 5** bir keşfin habercisidir!

Biraz ROOT

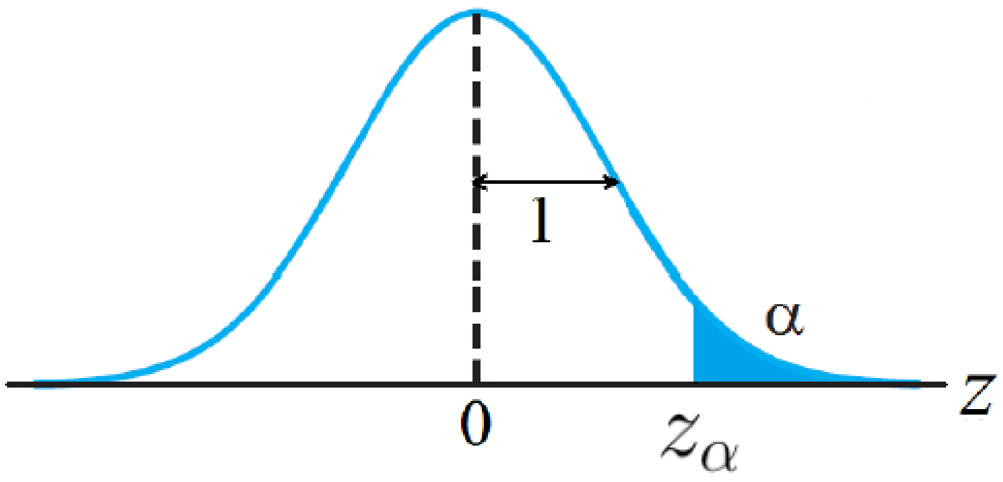
P-değeri: $p = P(Z > z_\alpha) = \alpha = \int_{z_\alpha}^{\infty} f(z)dz = \Phi(\infty) - \Phi(z_\alpha) = 1.0 - \Phi(z_\alpha)$

Önem: $Z = \Phi^{-1}(1 - p)$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

$\Phi(x) =$ **ROOT::Math::normal_cdf(x)**

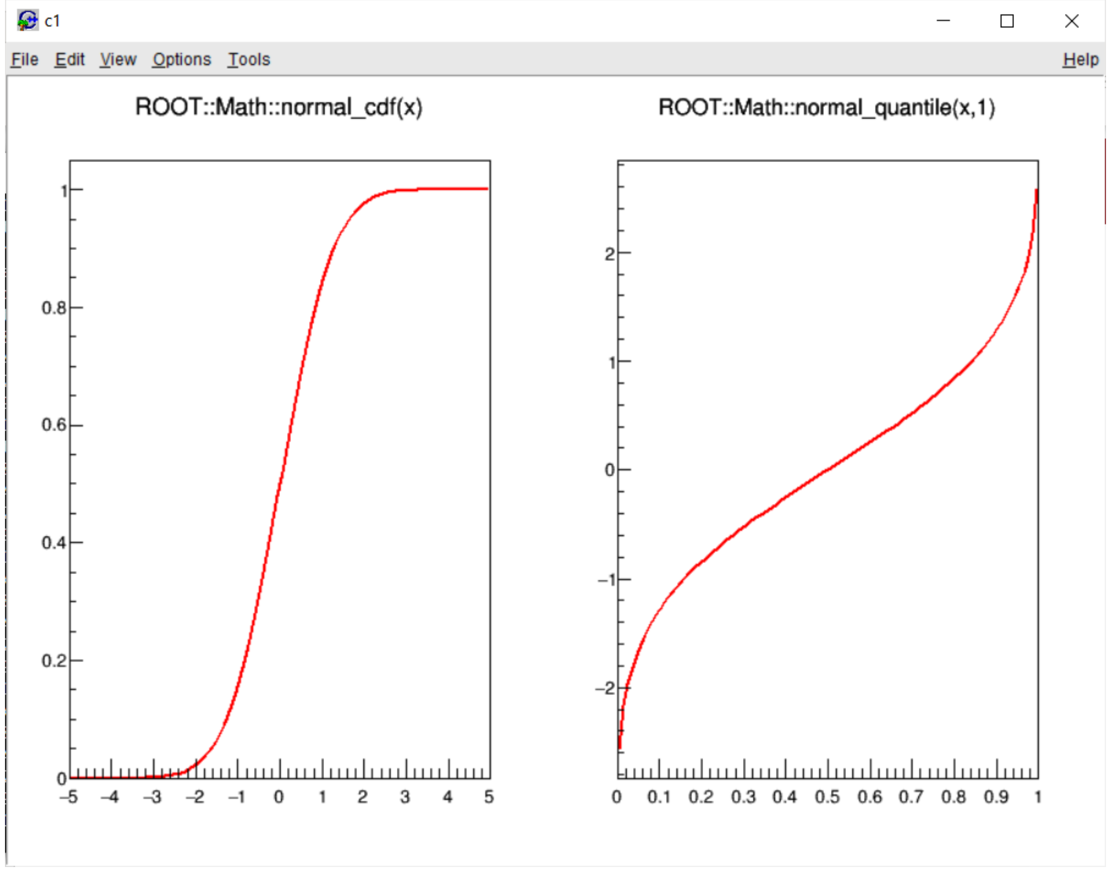
$\Phi^{-1}(x) =$ **ROOT::Math::normal_quantile(x,1)**



```
void norm(){
  TCanvas *c1 = new TCanvas("c1","c1",800,600);
  TF1 *phi     = new TF1("phi", "ROOT::Math::normal_cdf(x)",-5,5);
  TF1 *phi_inv= new TF1("phii","ROOT::Math::normal_quantile(x,1)",0,1);
  c1->Divide(2,1);
  c1->cd(1); phi->Draw();
  c1->cd(2); phi_inv->Draw();

  double alpha[5] = {0.1, 0.05, 0.01, 0.001, 0.0001};
  for(int i=0; i<5; i++){
    double  p = alpha[i];
    double CL = 1-p;
    cout << p << "\t" << CL <<"\t"
         << ROOT::Math::normal_quantile(CL,1) << endl;
  }
}
```

```
0.1     0.9     1.28155
0.05    0.95    1.64485
0.01    0.99    2.32635
0.001   0.999   3.09023
0.0001  0.9999  3.71902
```



Testlerin Uygulanması

Keşif problemlerinde:

H_0 sadece artalanı (background) tanımlar.

H_A sinyal + artalan ile belirlenir.

Eğer $P \leq 2.87\times10^{-7}$ ise, H_A kabul edilir ve veriyi daha iyi açıkladığı söylenir.

Dışarlama problemlerinde: (durum terstir)

H_0 sinyal + artalan ile belirlenir

H_A sadece artalan temsil eder.

H_0 hipotezi için test uygulanır ve üst limit konur. Genellikle $P \leq 0.05$ seçilir.

Gözlem ve Kuramın Karşılaştırılması

Ölçüm sonuçları kuramsal beklenen değerler ile karşılaştırıp, kuramlar test edilir. Fakat bu karşılaştırma için, deneysel ve kuramlar arasındaki farkların ölçüm belirsizliğinden ne kadar uzakta olduğu çok önemlidir..

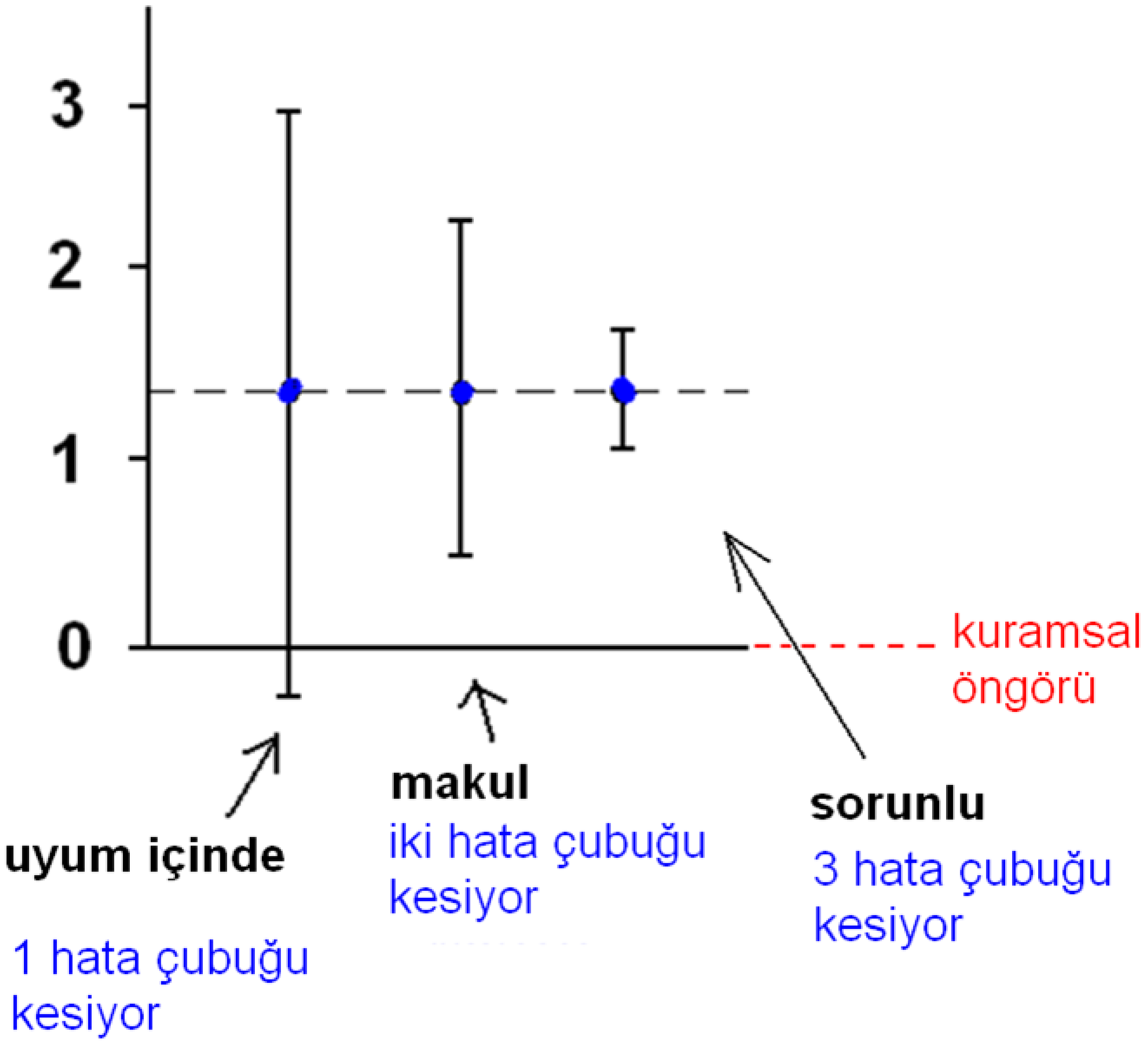
Örneğin kuramın öngördüğü değer 0.0 (sıfır) olsun.

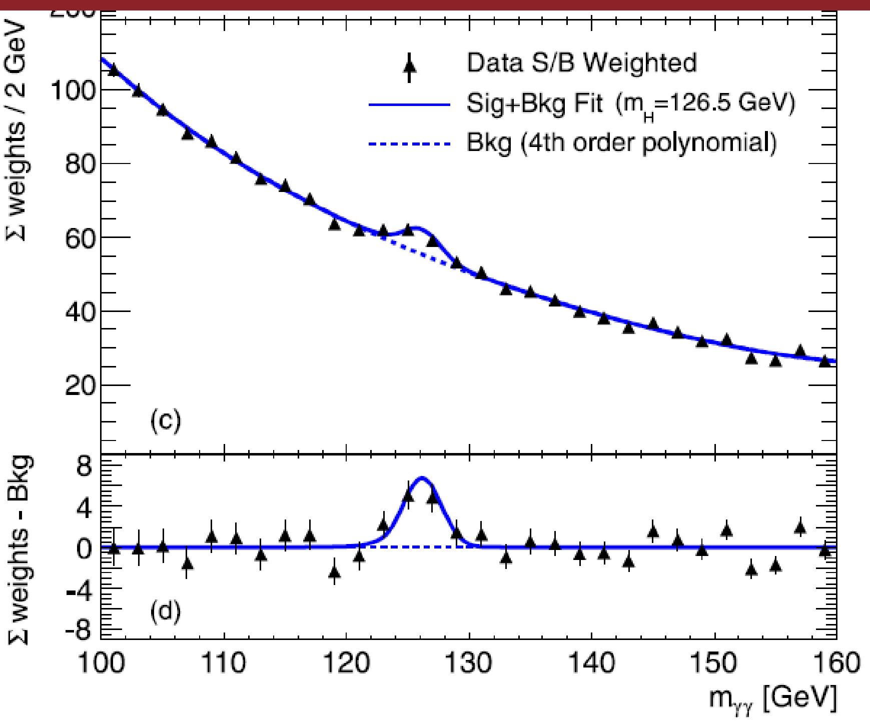
* Eğer ölçüm sonucu 1.3 ± 1.4 ise
  ölçüm ve kuram uyum içindedir. Çünkü Z = **(1.3-0.0)/1.4 = 0.93 < 1**

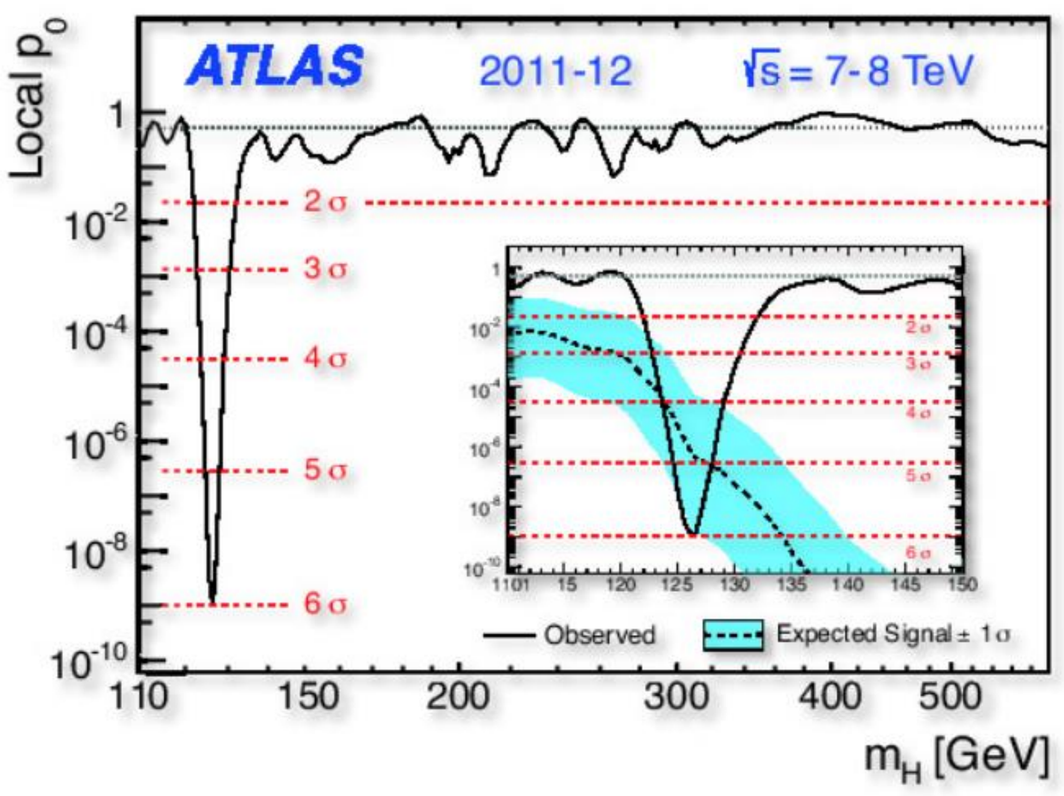
* Eğer ölçüm sonucu 1.3 ± 0.8 ise
  sonuç hala makuldür. Çünkü Z = **(1.3-0.0)/0.8 = 1.63 < 2**

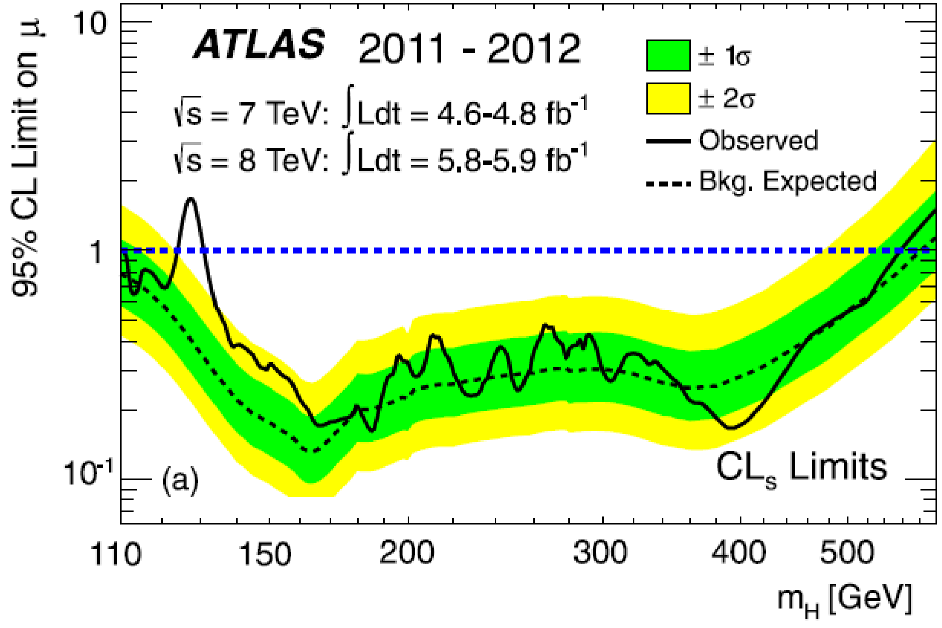
* Eğer ölçüm sonucu 1.3 ± 0.3 ise
  bu bir problemin işaretidir. Çünkü Z = **(1.3-0.0)/0.3 = 4.33 > 3**

Gözlem ve Kuramın Karşılaştırılması







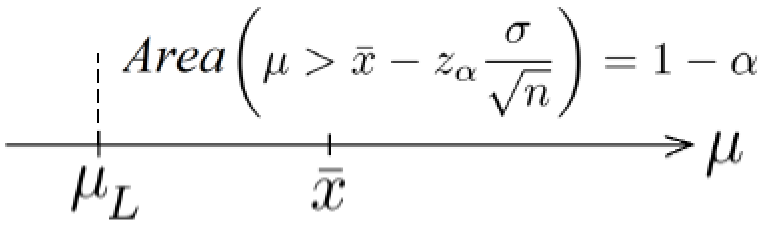


Her bir M_H değeri için sinyali bul ve CL (önemi) çiz.

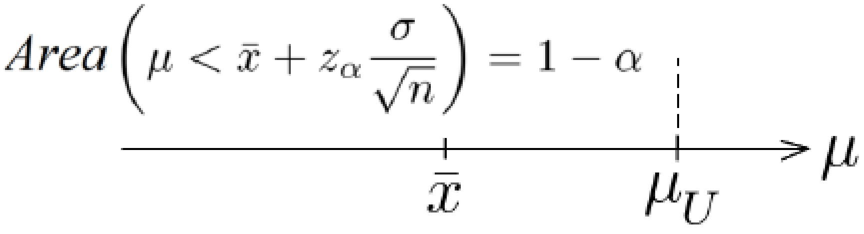
Küçük CL değerleri (büyük p değerleri), böyle bir artalandan sinyal bulma şansımızın az olduğunu gösterir.

Ayrıca, sinyal hipotezinin doğru olduğu durumda beklenen önem (MC) da beraber çizdirilebilir.

Tek Taraflı Limit Koyma



Lower one-sided bound on μ.



Upper one-sided bound on μ.

Örnek: Nötrino kütlesi ölçme deneyinde 300 adet $\pi^+ \to \mu^+ + \nu_\mu$ bozunumu incelenmiş ve nötrinonun kütlesi $m_\nu = 0.18 \pm 0.27$ MeV olarak tespit edilmiştir. Buradaki hata standart sapmadır. Nötrino kütlesinin %95 CL düzeyinde üst limitini belirleyin.

$\bar{x} = 0.18\ MeV$

$\sigma = 0.27\ MeV$

$n = 300$

$$e = z_\alpha \frac{\sigma}{\sqrt{n}} = 1.645 \frac{0.27}{\sqrt{300}} = 0.026\ \text{MeV}$$

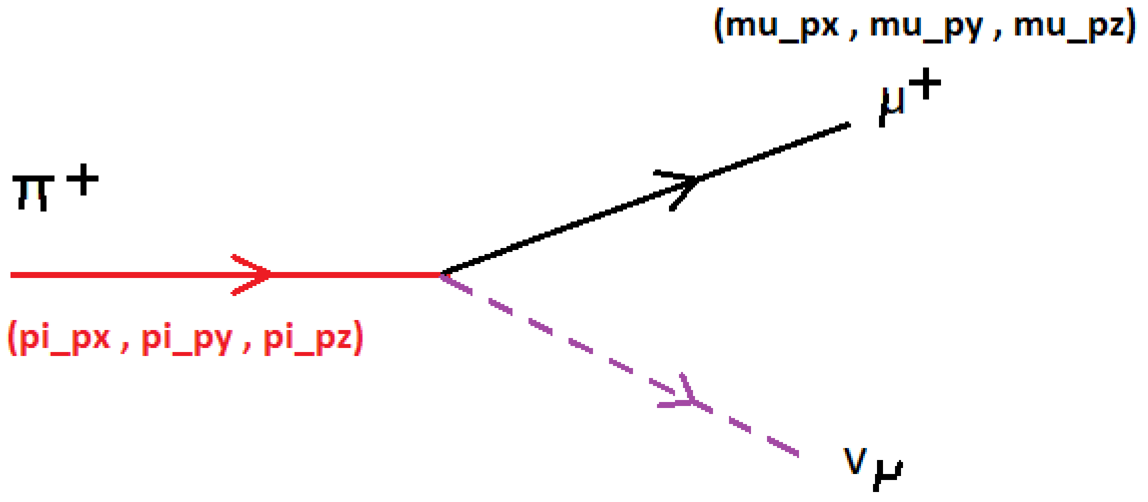
Buna göre, üst limit:

$m_\nu < 0.180 + 0.026$ => $\boldsymbol{m_\mu < 0.21\ \text{MeV}}$ **CL = 95%**

We are 95% confident that the mean neutrino mass is less than 0.21 MeV.

Alıştırma

Ekte sunulan root programı, NeutrinoMass.C, her olayda $\pi^+ \rightarrow \mu^+ + \nu_\mu$ bozumununu Monte Carlo simulayonunu gerçekleştirir.



Program m_ν^2 ve her parçacığın momentum dağlımlarını çizer. Nötrino kütlesinin %90 düzeyinde üst limitini hesaplar.

Not: Hata yayılımı (error propagation)

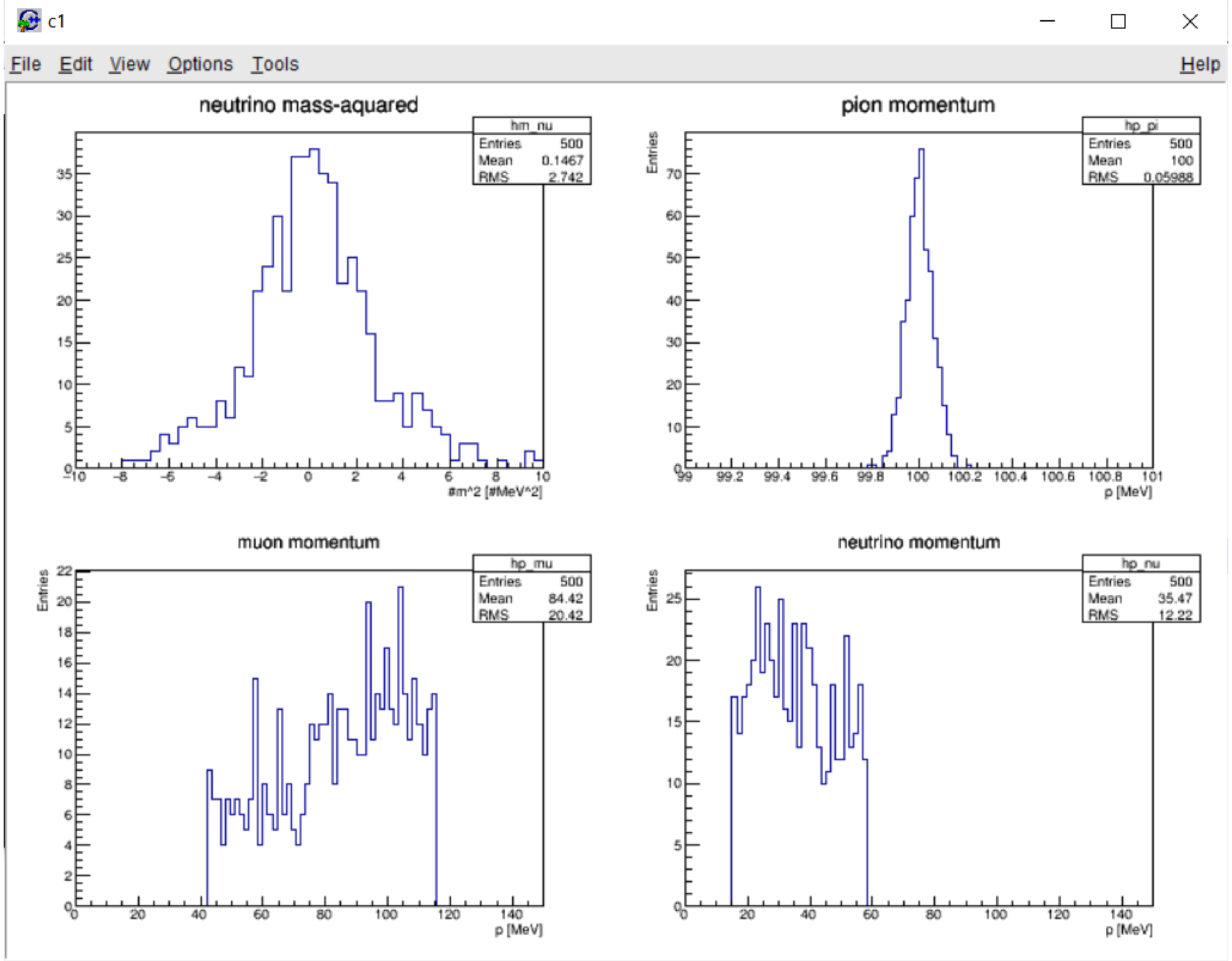
$f(x) = x^2$ *fonksiyondur. Fonksiyondaki belirsizlik* σ_f *ise,* x *değerinin belirsizliği*

$\sigma_f = \frac{df}{dx}\sigma_x = 2x\ \sigma_x$ *ile hesaplanır.*

Alıştırma

Programın Çıktısı:

```
Mean of neutrino mass-square = 0.15
RMS  of neutrino mass-square = 2.7
Mass neutrino < 0.59 MeV  CL = 90%
```



KAYNAKLAR

[1]. https://pdg.lbl.gov/2021/reviews/rpp2020-rev-passage-particles-matter.pdf (2019)

[2]. W.R. Leo, *Techniques for Nuclear and Particle Physics Experiments 2nd Ed*, Springer (1994)