Phys 491 Fall 2021

Lecture 2: Some Probability

1. Histogram graph of frequency distribution

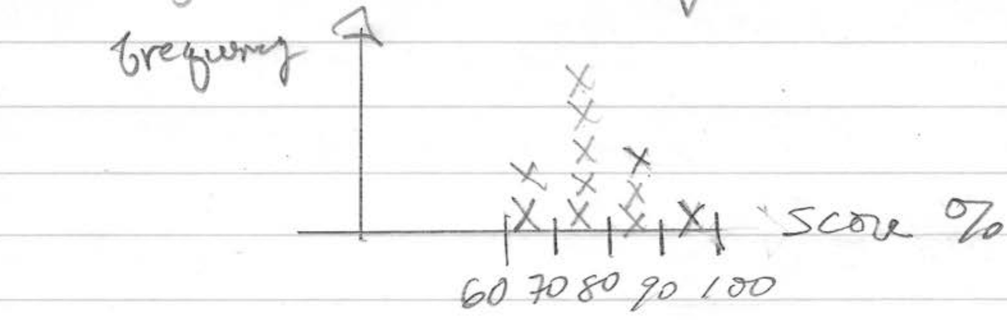
Consider grades on an exam: N students, $\{g_i\}$

mean $\langle g \rangle = \frac{1}{N} \sum g_i$

Variance $V = (\Delta g)^2 = \frac{1}{N} \sum (g_i - \langle g \rangle)^2$
= square of uncertainty

Note - for unbiased estimate of variance, use $N-1$

Grade distribution visualized by binning data and making histogram



bin	Center X_i	number N_i
7	60.5	2
8	70.5	5
9	80.5	3
10	90.5	1

$\langle x \rangle = \frac{1}{N} \sum N_i X_i \qquad (\Delta x)^2 = \frac{1}{N} \sum N_i (X_i - \langle x \rangle)^2$

Imagine there is some underlying true probability distribution. We can use data to estimate

$$p_i \approx \frac{N_i}{N} \qquad \boxed{\sum P_i = 1}$$

$$\langle x \rangle = \sum P_i x_i$$

$$\langle x^2 \rangle = \sum P_i x_i^2$$

$$\langle (x - \langle x \rangle)^2 \rangle = \sum P_i (x - \langle x \rangle)^2$$

$$= \sum P_i x^2 - 2\langle x \rangle \sum P_i x + \langle x \rangle^2 \sum P_i$$

$$= \langle x^2 \rangle - \langle x \rangle^2$$

In continuous limit:

$$P_i \longrightarrow p(x)\,dx$$

$$\langle x \rangle = \int_{-\infty}^{\infty} P(x)\, x\, dx$$

P(x) is a probability density function (PDF) where

$$\int_{-\infty}^{\infty} p(x)\,dx = 1$$

Example: exponential decay

$$\frac{dN}{dt} = -\frac{N(t)}{\tau} \Rightarrow N(t) = N_0 e^{-t/\tau}$$
$$N_0 \equiv N(0)$$

normalized PDF is $p(t) = 0,\ t < 0$
$$p(t) = \frac{1}{\tau} e^{-t/\tau}$$

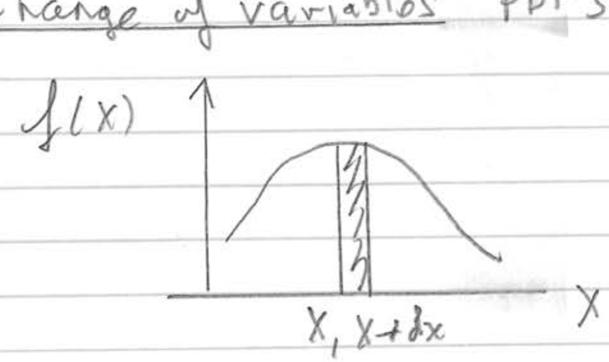
$$\langle t \rangle = \frac{1}{\tau}\int_0^\infty e^{-t/\tau}\, t\, dt = \tau \int_0^\infty e^{-x} x\, dx = \tau$$

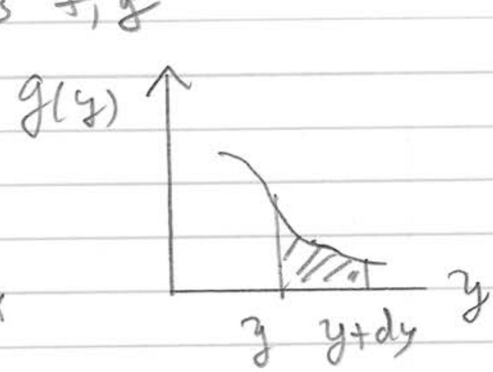
useful integral $\int_0^\infty x^n e^{-x}\, dx = n!$

Variance $(\Delta t)^2 = \frac{1}{\tau}\int e^{-t/\tau} (t-\tau)^2\, dt$
$$= \tau^2 \int_0^\infty e^{-x}(x-1)^2\, dx = \tau^2 [2! - 2 + 1] = \tau^2$$

$$\Delta t = \tau$$

Change of variables PDF's f, g





change of variables $y = y(x)$ inverted by $x = x(y)$

$$f(x)\,dx = g(y)\,dy$$

$$g(y)\,dy = \left|\frac{dy}{dx}\right|^{-1} dy\, f\big(x(y)\big)$$

so PDF $g(y) = \left|\frac{dy}{dx}\right|^{-1} f\big(x(y)\big)$

Example

$$f(x) = e^{-x} \qquad y = e^{-x}$$

$$x(y) = -\ln(y) \qquad \left|\frac{dy}{dx}\right| = e^{-x} = y$$

$$g(y) = \left|\frac{dy}{dx}\right|^{-1} f\big(x(y)\big)$$

$$= \frac{1}{y}\, y = 1 \quad \text{flat}$$

range:
$x\ [0, \infty] \qquad y\ [0, 1]$

Error Propagation

measure two observables x, y to determine $q(x,y)$
Generally, there is a PDF in 2 variable

$$\int P(x,y)\, dx\, dy = 1$$

$$\langle x \rangle = \int P(x,y)\, x\, dx\, dy \quad , \text{etc.}$$

$$\langle q \rangle = \int P(x,y)\, q(x,y)\, dx\, dy$$

expand $q(x,y)$ about $\langle x \rangle, \langle y \rangle$

$$q(x,y) = q(\langle x\rangle, \langle y \rangle) + \frac{\partial q}{\partial x}\Big| (x - \langle x \rangle) + \frac{\partial q}{\partial y}\Big| (y - \langle y \rangle)$$

where $\frac{\partial q}{\partial x}\Big|, \frac{\partial q}{\partial y}\Big|$ are evaluated at $\langle x \rangle, \langle y \rangle$

and are not function of x, y.

then $\langle q \rangle = q(\langle x \rangle, \langle y \rangle)$ and $\langle x - \langle x \rangle \rangle = 0 = \langle y - \langle y \rangle \rangle$

$$V_q = (\Delta q)^2 = \langle (q - \langle q \rangle)^2 \rangle =$$

$$\left(\frac{\partial q}{\partial x}\Big|\right)^2 \langle (x - \langle x \rangle)^2 \rangle + \left(\frac{\partial q}{\partial y}\Big|\right)^2 \langle (y - \langle y \rangle)^2 \rangle$$

$$+ 2 \frac{\partial q}{\partial x}\Big| \frac{\partial q}{\partial y}\Big| \underbrace{\langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle}_{= \mathrm{Cov}(x,y)} \quad \text{covariance}$$

giving

$$(\Delta q)^2 = \left(\frac{\partial q}{\partial x}\Big|\right)^2 (\Delta x)^2 + \left(\frac{\partial q}{\partial y}\Big|\right)^2 (\Delta y)^2$$

$$+ \; 2 \frac{\partial q}{\partial x}\Big| \frac{\partial q}{\partial y}\Big| \; \mathrm{Cov}(x, y)$$

if x, y are uncorrelated $P(x,y) = P_x(x) P_y(y)$

and $\mathrm{Cov}(x, y) = 0$

linear correlation coefficient

$$r \equiv \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\Delta x \, \Delta y} = \frac{\mathrm{Cov}(x, y)}{\Delta x \, \Delta y}$$

$-1 \leq r \leq 1$ linearly independent

$|r| = 1$ perfect linear correlation

Important probability distributions

Binomial Two outcomes with probabilities p, $q = 1-p$.

Binomial theorem

$$(p+q)^N = 1 = \sum_{r=0}^{N} \binom{N}{r} p^r q^{N-r}$$

binomial coefficient symbol

$$\binom{N}{r} \equiv \frac{N!}{r!\,(N-r)!}$$

ways of selecting r in any order out of N
recall $0! = 1$ (1 way to arrange zero objects, null set)

$$\langle r \rangle = pN \quad \text{or} \quad \left\langle \frac{r}{N} \right\rangle = p$$

proof: $\langle r \rangle = \sum_{r=0}^{N} \binom{N}{r} r\, p^r q^{N-r}$

$$\frac{d}{dp}(p+q)^N = 0 = \sum_{r=0}^{N} \binom{N}{r} \left[r p^{r-1} q^{N-r} + (N-r) p^r q^{N-r-1} \overset{-1}{\overset{''}{\frac{dq}{dp}}} \right]$$

multiply by p

$$\langle r \rangle = p \sum_{r=0}^{N} (N-r) \frac{N!}{r!\,(N-r)!} p^r q^{N-r-1}$$

$r=N$ term is zero

$$= Np \sum_{r=0}^{N-1} \frac{(N-1)!}{r!\,(N-r-1)!} p^r q^{N-1-r} = Np(p+q)^{N-1} = Np$$

In a similar way you can prove

$$V(r) = (\Delta r)^2 = Np(1-p)$$

$$\Delta\left(\frac{r}{N}\right) = \sqrt{\frac{p(1-p)}{N}}$$

In data analysis suppose you get
N_1 outcomes with two possibilities in N trials.

$p \approx \frac{N_1}{N}$ with estimated error

$$\Delta p = \sqrt{\frac{N_1(N-N_1)}{N^3}}$$

Poisson limit of binomial $p \to 0$, $N \to \infty$
with $\mu = pN$ constant

Probability to have r events in fixed (space or time) interval assuming events are independent with mean μ

$$P(r,\mu) = \frac{\mu^r}{r!} e^{-\mu}$$

For example, suppose 2 people per 10 minutes on average enter a store. Prob for 5 people to enter store in 10 minutes is

$$P(5,2) = \frac{2^5}{5!} e^{-2} = 0.036 \quad P_{\geq 5} = 0.053$$

<u>Gaussian</u> Extremely important due to central limit theorem

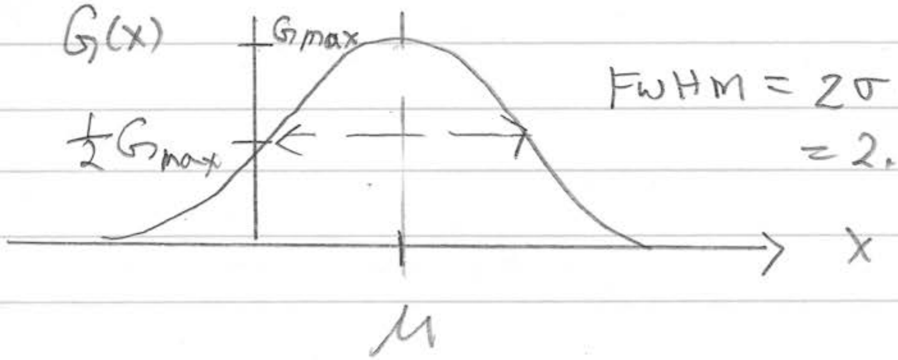
limit of Binomial as $N \to \infty$

$$G(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}$$

parameters

$\langle x \rangle_G = \mu \qquad (\Delta x)_G = \sigma$

σ parameter is Gaussian error. Often we refer to variance as $V = \sigma^2$ even for non-Gaussian distributions



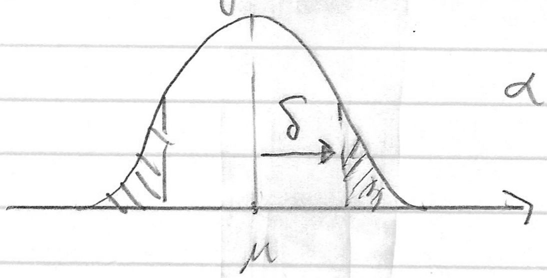
$\text{FWHM} = 2\sigma\sqrt{2\ln 2}$
$= 2.35\,\sigma$

"Bell Curve."

A <u>parabola</u> on semi-log plot

$$\log G = -\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2$$

Probability content of Gaussian



α is hatched area

α	δ
37.13 %	σ
4.55 %	2σ
0.27%	3σ
6.3×10^{-5}	4σ

α	δ
20%	1.28σ
10%	1.64σ
5%	1.96σ
1%	2.58σ

Read Appendix D in Townsend to learn how to do Gaussian integrals.

Intermediate Quantum 491: Central Limit Theorem

- Theorem: Given N independent random variables r_i each distributed according to PDFs R_i having finite mean μ_i and variance V_i,

$$y = \frac{\sum_{i=1}^{N} r_i}{N}$$

tends towards a Gaussian, and becomes Gaussian with parameters μ, σ in the limit $N \to \infty$.

$$\mu = \langle \mathrm{y} \rangle = \lim_{\mathrm{N}\to\infty} \frac{\sum_{\mathrm{i}=1}^{\mathrm{N}} \mu_{\mathrm{i}}}{\mathrm{N}}$$

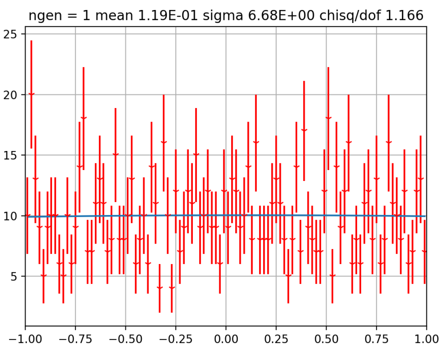
$$\sigma^2 = (\Delta \mathrm{y})^2 = \mathrm{V_y} = \lim_{\mathrm{N}\to\infty} \frac{\sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{V_i}}{\mathrm{N}}$$

- Numerical Test

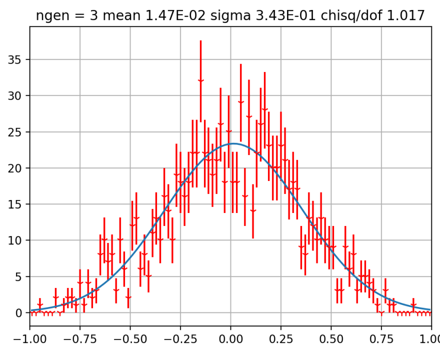
A simple example shows how rapidly the Gaussian is reached. Take all R_i as the same uniform random numbers with mean 0.

  - Generate N uniform random numbers in the interval $[-1, 1]$.
  - Find the average of the N numbers
  - Repeat 1000 times.

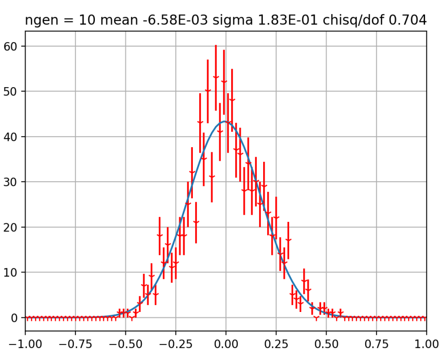
Average of N uncorrelated numbers rapidly becomes Gaussian with increasing N, with width of Gaussian shrinking as $1/\sqrt{N}$.


ngen = 1 mean 1.19E-01 sigma 6.68E+00 chisq/dof 1.166

(a) N=1


ngen = 3 mean 1.47E-02 sigma 3.43E-01 chisq/dof 1.017

(b) N=3


ngen = 10 mean -6.58E-03 sigma 1.83E-01 chisq/dof 0.704

(a) N=10