Lecture 11

Central Limit Theorem:

average of almost any type of RV, in the limit of large N (# of trials), will be Gaussian distributed.

X is an r.v.

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

sample mean

$$\bar{X} \sim G(\mu, \sigma^2)$$, Gaussian R.V.

$\mu \equiv$ mean

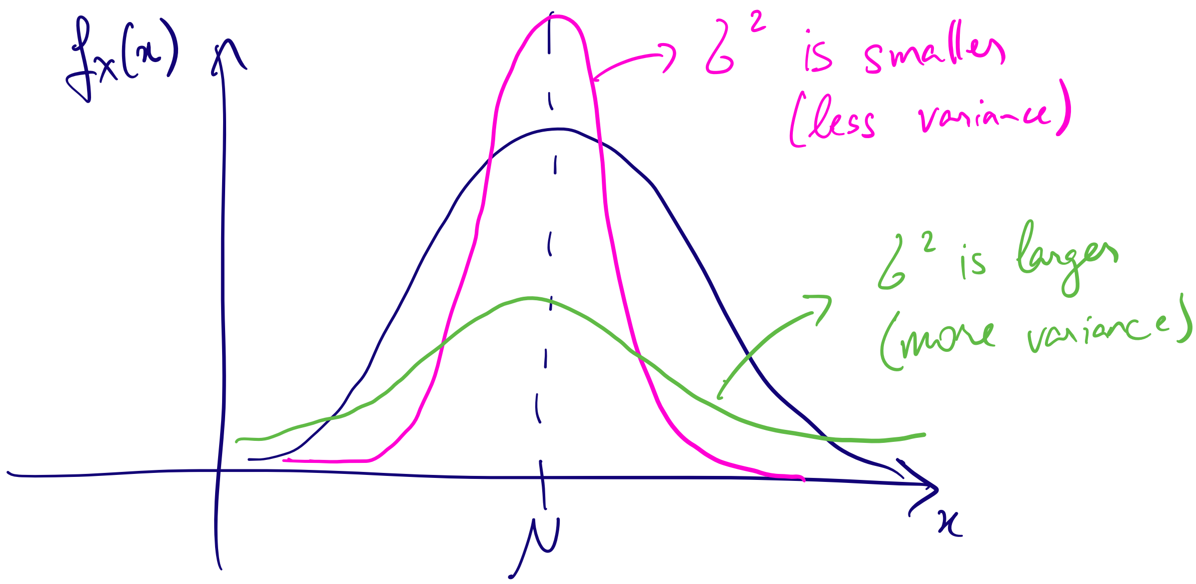
$\sigma^2 \equiv$ variance

$\sigma \equiv$ standard deviation

$$X \sim G(\mu, \sigma^2)$$

PDF:

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



Gaussian PDF is symmetric w.r.t. the mean value μ.

with respect to.

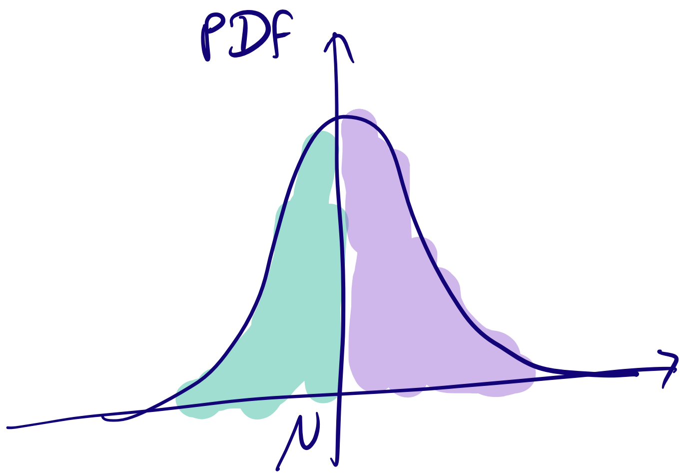
Symmetric $\Rightarrow$ mean value is equal to the median value

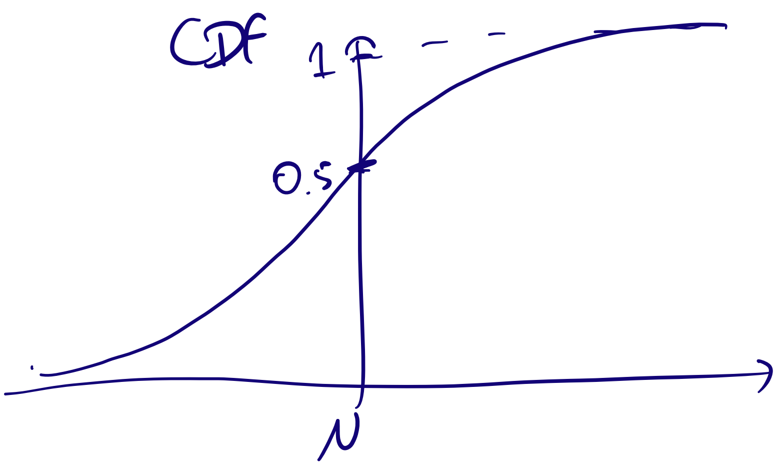
$$X \sim G(0, 1)$$

μ σ^2

Standard Normal distribution

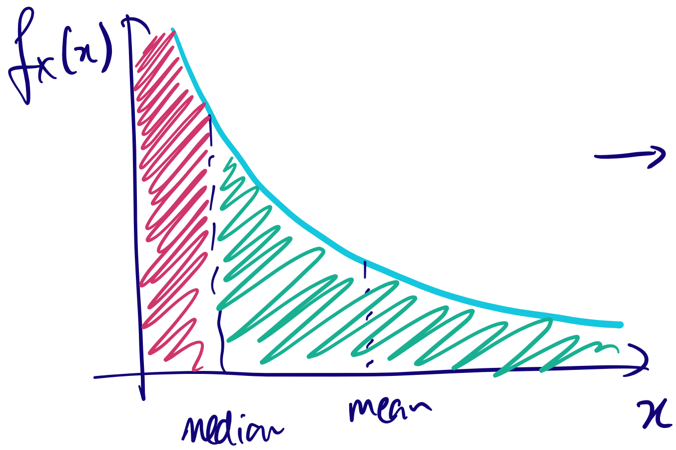
$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$



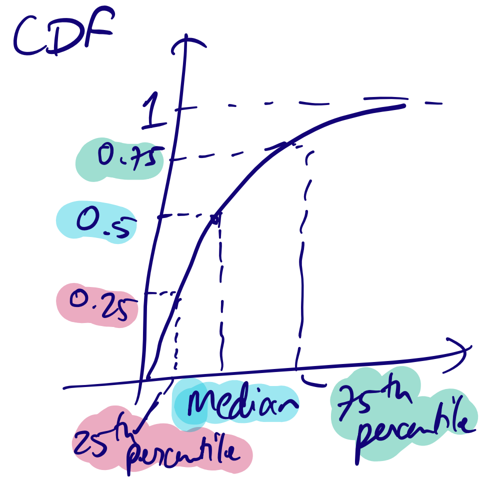


Non-symmetric Distribution : median $\neq$ mean

Exponential R.V.



$\rightarrow$



median : x such that $F_X(x) = 0.5$

25th percentile : x such that $F_X(x) = 0.25$

75th " : x " " " $= 0.75$

$$x_{75th} = F_X^{-1}(0.75)$$

Properties of PDFs (Prob. Density Functions)

① $F_X(x) = \int_{-\infty}^{x} f_X(u)\, du$

Proof: Fundamental Theorem of Calculus.

$$f_X(x) = \frac{dF_X(x)}{dx}$$

② $f_X(x) \geq 0 \quad , \quad -\infty < x \leq \infty$

Proof: $F_X(x)$ is monotonically non-decreasing

$$f_X(x) = \frac{dF_X(x)}{dx}$$

$$\Downarrow$$

$$f_X(x) \geq 0$$

③ $\int_{-\infty}^{\infty} f_X(x) \cdot dx = 1$

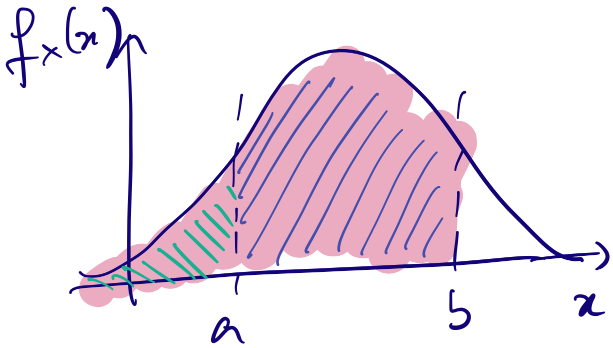
If $f_X(x)$ satisfies this property, we say that it's a valid PDF!

(4) $P(a < x \leq b) = \int_a^b f_x(x)\,dx$

where $a \leq b$

Proof: $P(a < x \leq b) = F_x(b) - F_x(a)$

$= P(X \leq b) - P(X \leq a)$



$= \int_{-\infty}^{b} f_x(x)\,dx - \int_{-\infty}^{a} f_x(x)\,dx$

$= \int_a^b f_x(x)\,dx$

(5) If $g(x)$ is a non-negative continuous function w/ a finite integral,

$\int_{-\infty}^{\infty} g(x)\,dx = c$, where c is a constant $c \neq 0$

Then $f_x(x) = \dfrac{g(x)}{c}$ is a valid PDF.

Gaussian R.V.s :

CDF: $F_X(x) = P(X \leq x)$

$$= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt$$

Standard Normal Gaussian R.V.

$\mu = 0$
$\sigma^2 = 1$

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-t^2}{2}\right) dt$$

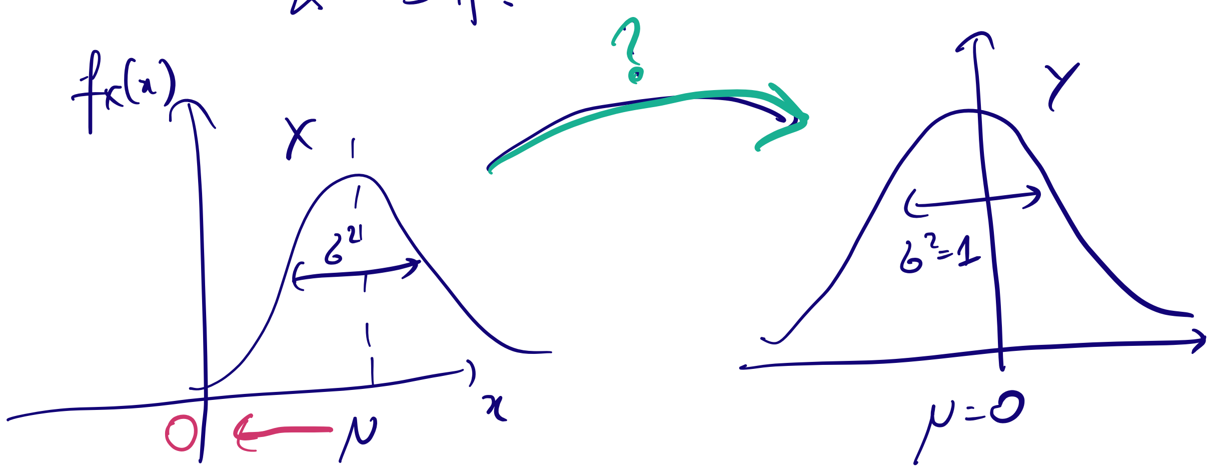
$Q(x) = 1 - \Phi(x)$ survival function of a Gaussian(0,1)

Q-function

For value of μ and σ^2, we have a different CDF!

1. Scale Gaussian (μ, σ^2) to be a standard normal $G(0,1)$

2. Use the lookup table (Q-function table) to find the values for its CDF & s.f.



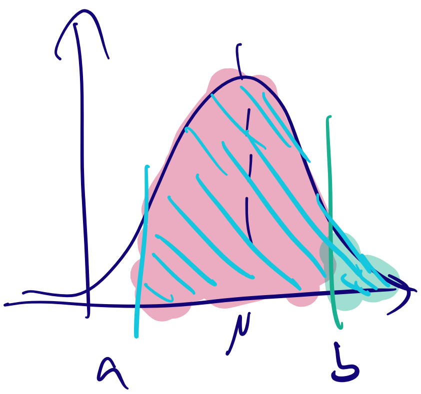
$$\boxed{Y = \frac{X-\mu}{\sigma}}$$

$X \sim G(\mu, \sigma^2) \longrightarrow Y \sim G(0,1)$

② $Q(x) = 1 - \Phi(x)$ of Gaussian $(0,1)$

$X \sim G(\mu, \sigma^2)$

$P(\underline{a < X} \leq b) = \overset{①}{P(X > a)} - \overset{②}{P(X > b)}$



① $P\left(\underbrace{\frac{X-\mu}{\sigma}}_{G(0,1)} > \frac{a-\mu}{\sigma}\right)$

$P\left(Y > \frac{a-\mu}{\sigma}\right) = Q\left(\frac{a-\mu}{\sigma}\right)$

Q-function : $P(Y > y)$

② $P(X > b)$

$= P\left(\frac{X-\mu}{\sigma} > \frac{b-\mu}{\sigma}\right)$

$= P\left(\underset{\sim G(0,1)}{Y} > \frac{b-\mu}{\sigma}\right)$

$= Q\left(\frac{b-\mu}{\sigma}\right)$

$\therefore\ P(a < x \leq b) = Q\left(\frac{a-\mu}{\sigma}\right) - Q\left(\frac{b-\mu}{\sigma}\right)$

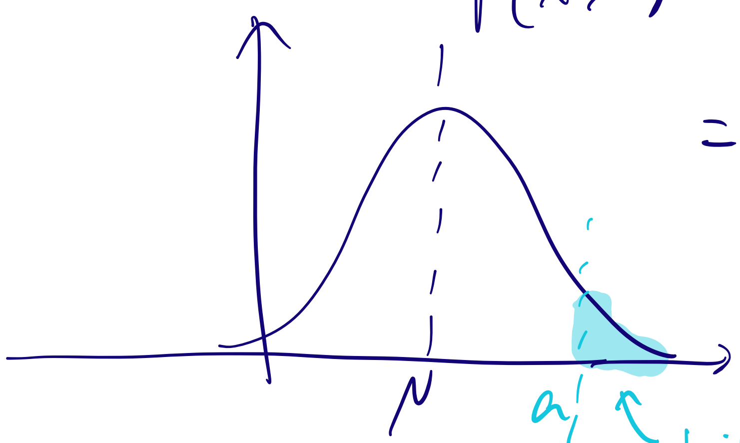
$\sim G(\mu, \sigma^2)$

$X \sim G(\mu, \sigma^2)$

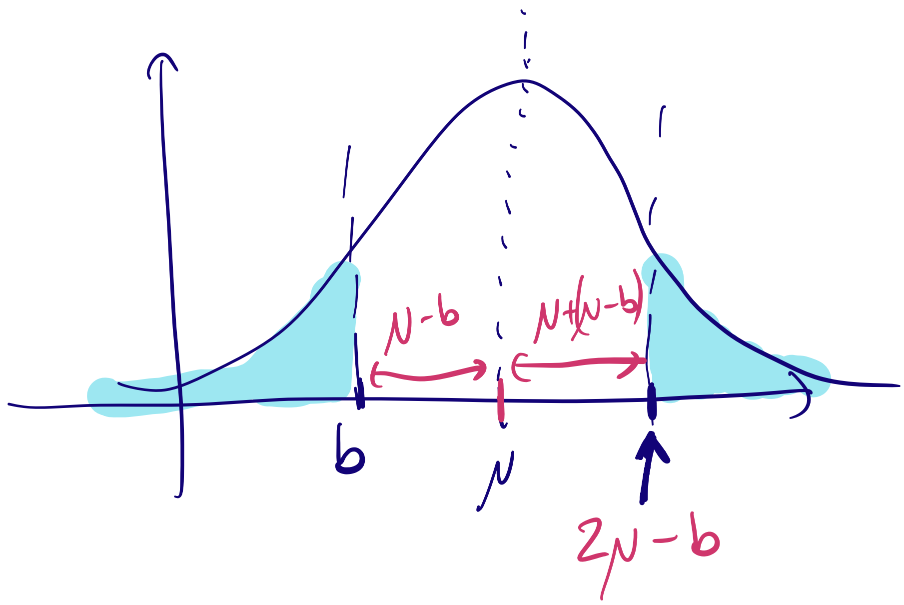
<u>Case 1</u>: $P(X \geq a)$ where $a \geq \mu$:

$$P(X \geq a) = P(X > a)$$

$$= Q\left(\frac{a-\mu}{\sigma}\right)$$



Case 2 : $P(X \le b)$ where $b < \mu$



$$P(x \le b) = P(x > 2\mu - b)$$

$$= Q\left(\frac{(2\mu - b) - \mu}{\sigma}\right)$$

$$= Q\left(\frac{\mu - b}{\sigma}\right) = Q\left(\frac{-(b - \mu)}{\sigma}\right)$$

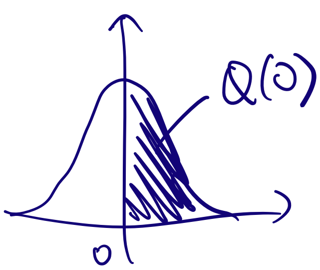
Example: Grading on a curve.

$X \equiv$ R.V. represents a student's grade.

$X \sim G(\mu, \sigma^2)$

a) What is the probability that grade is above μ?

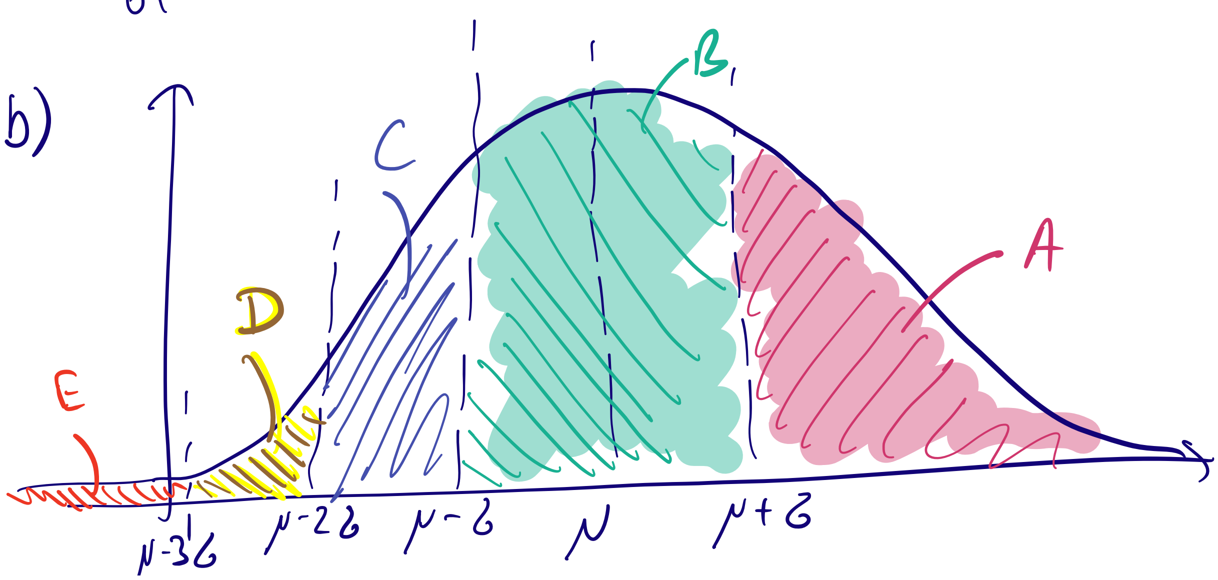
$$\Rightarrow P(X > \mu) = P\Big(\overbrace{\frac{X-\mu}{\sigma}}^{Y \sim G(0,1)} > \frac{\mu-\mu}{\sigma}\Big)$$



$$= P(Y > 0)$$

$$= Q(0) = 1/2$$

b)



$$P(A) = P(X \geq \mu + \sigma) = Q\left(\frac{\mu + \sigma - \mu}{\sigma}\right) = Q(1)$$

$$P(B) = P(\mu - \sigma < X < \mu + \sigma)$$

$$= 1 - 2\,P(X > \mu + \sigma) \quad \text{by symmetry}$$

$$= 1 - 2\,Q\left(\frac{\mu + \sigma - \mu}{\sigma}\right)$$

$$= 1 - 2Q(1)$$

$$P(C) = P(\mu - 2\sigma < X < \mu - \sigma)$$

$$= P(X > \mu - 2\sigma) - P(X > \mu - \sigma)$$

$$= Q\left(\frac{\mu - 2\sigma - \mu}{\sigma}\right) - Q\left(\frac{\mu - \sigma - \mu}{\sigma}\right)$$

symmetry

$$= Q(-2) - Q(-1)$$

$$= (1 - Q(2)) - (1 - Q(1))$$

$$= Q(1) - Q(2)$$

$$Q(-a) = P(x > -a) = 1 - P(x > a)$$
$$= 1 - Q(a)$$

For any value of a:

$$Q(-a) = 1 - Q(a)$$

$$P(D) = P(\mu - 3\sigma < x < \mu - 2\sigma)$$
$$= P(x > \mu - 3\sigma) - P(x > \mu - 2\sigma)$$
$$\vdots$$
$$= Q(2) - Q(3)$$

$$P(E) = P(x \leq \mu - 3\sigma)$$

by symmetry

$$= P(x > \mu + 3\sigma)$$
$$= Q\left(\frac{\mu + 3\sigma - \mu}{\sigma}\right) = Q(3)$$

c) $P(x > \mu + k\sigma) = 0.4$

$$\Leftrightarrow Q\left(\frac{\mu + k\sigma - \mu}{\sigma}\right) = 0.4$$

$\Leftrightarrow \quad Q(k) = 0.4$

$\Leftrightarrow \quad k = Q^{-1}(0.4) \Rightarrow k \simeq 0.25$

$\boxed{\text{stats.norm}}$

$\boxed{\text{stats.norm.sf}(x)} \Rightarrow Q(x)$!!

Gaussian(0,1)

$\boxed{\text{stats.norm.isf}(x)}$