COMP 1433: Introduction to Data Analytics &
COMP 1003: Statistical Tools and Applications

Tutorial 3 – Statistic Basics

Chuan He
Department of Computing
The Hong Kong Polytechnic University

Proof

1. For two random variables X and Y, prove $E(X+Y) = E(X) + E(Y)$

2. If X and Y are ***independent*** random variables, then prove $E(XY) = E(X) * E(Y)$

3. For random variable X, prove $Var(X) = E(X^2) - E(X)^2$

4. If X and Y are ***independent*** random variables, then prove $Var(X+Y) = Var(X) + Var(Y)$

Expectation and Application

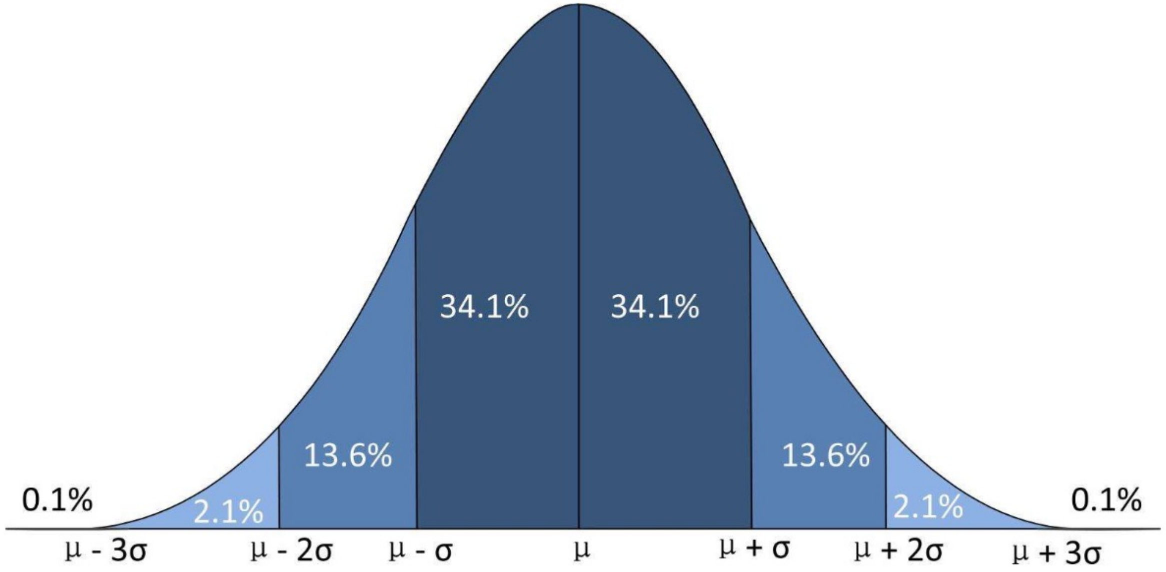
A certain kind of lizard lays 8 eggs, each of which will hatch independently with probability 0.7. Let Y denote the number of eggs which hatch. Then $Y \sim B(8, 0.7)$ (binomial distribution) . What is the PMF, CDF, expectation and variance of Y?

Normal Distribution

Let X have the standard normal distribution with density

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

Find the density of $Y = \sigma X + \mu$ for given constants μ and $\sigma \neq 0$. Also, find the density of $Z = X^2$.



Proof

1. For two random variables X and Y, prove $E(X+Y) = E(X) + E(Y)$

2. If X and Y are ***independent*** random variables, then prove $E(XY) = E(X) * E(Y)$

3. For random variable X, prove $\mathrm{Var}(X) = E(X^2) - E(X)^2$

4. If X and Y are ***independent*** random variables, then prove $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

Expectation and Application

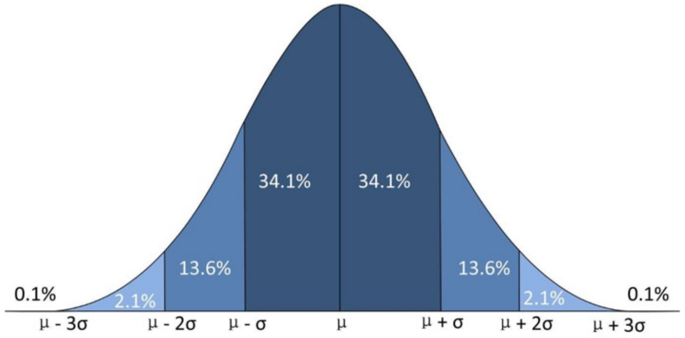
A certain kind of lizard lays 8 eggs, each of which will hatch independently with probability 0.7. Let Y denote the number of eggs which hatch. Then $Y \sim B(8, 0.7)$ (binomial distribution) . What is the PMF, CDF, expectation and variance of Y?

Normal Distribution

Let X have the standard normal distribution with density

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

Find the density of $Y = \sigma X + \mu$ for given constants μ and $\sigma \neq 0$. Also, find the density of $Z = X^2$.



Time for exercises

Proof

1. For two random variables X and Y, prove $\mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y)$.

2. If X and Y are ***independent*** random variables, then prove $\mathrm{E}(XY) = \mathrm{E}(X) * \mathrm{E}(Y)$

3. For random variable X, prove $\mathrm{var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2$

4. If X and Y are ***independent*** random variables, then prove $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

Proof

1. For two random variables *X* and *Y*, prove E(*X* +*Y*) = E(*X*) + E(*Y*). (***always***)

$$S_{X+Y} = \{x+y \mid (x \in S_X) \cap (y \in S_Y)\},$$

$$\mathrm{E}(X+Y) = \sum_{(x+y)\in S_{X+Y}} (x+y)\,\mathrm{P}((X=x)\cap(Y=y)) = \sum_{x\in S_X}\sum_{y\in S_Y} (x+y)\,\mathrm{P}((X=x)\cap(Y=y))$$

$$= \sum_{x\in S_X}\sum_{y\in S_Y} x\,\mathrm{P}((X=x)\cap(Y=y)) + \sum_{x\in S_X}\sum_{y\in S_Y} y\,\mathrm{P}((X=x)\cap(Y=y))$$

P(XY) = P(X)P(Y|X)

$$\boxed{= \sum_{x\in S_X}\sum_{y\in S_Y} x\,\mathrm{P}(X=x)\;\mathrm{P}(Y=y|X=x)}\;\boxed{+ \sum_{y\in S_Y}\sum_{x\in S_X} y\,\mathrm{P}(Y=y)\;\mathrm{P}(X=x|Y=y)}$$

$$= \sum_{x\in S_X} x\,\mathrm{P}(X=x) \sum_{y\in S_Y} \mathrm{P}(Y=y|X=x) + \sum_{y\in S_Y} y\,\mathrm{P}(Y=y) \sum_{x\in S_X} \mathrm{P}(X=x|Y=y)$$

$$= \sum_{x\in S_X} x\,\mathrm{P}(X=x) + \sum_{y\in S_Y} y\,\mathrm{P}(Y=y)$$

$$= \mathrm{E}(X) + \mathrm{E}(Y).$$

Proof

2. If *X* and *Y* are ***independent*** random variables, then prove E(*XY*) = E(*X*) * E(*Y*).

$$S_{XY} = \{xy \mid (x \in S_X) \cap (y \in S_Y)\},$$

$$\begin{aligned}
\mathrm{E}(XY) &= \sum_{xy \in S_{XY}} xy\, \mathrm{P}((X = x) \cap (Y = y)) \\
&\boxed{= \sum_{x \in S_X} \sum_{y \in S_Y} xy\, \mathrm{P}(X = x)\ \mathrm{P}(Y = y)} \\
&= \sum_{x \in S_X} x\, \mathrm{P}(X = x) \sum_{y \in S_Y} y\, \mathrm{P}(Y = y) \\
&= \mathrm{E}(X)\ \mathrm{E}(Y)
\end{aligned}$$

Proof

3. For random variable X, prove Var(X) = E(X^2) – E(X)^2

$$Var(X) = E[(X-\mu)^2]$$

$$\begin{aligned} E\left(X^2\right) &= E\left((X-\mu+\mu)^2\right) \\ &= E\left((x-\mu)^2+\mu^2+2\mu(X-\mu)\right) \\ &= E\left((x-\mu)^2\right)+E\left(\mu^2\right)+2\mu E(X-\mu) \\ &= \operatorname{Var}(X)+E(X)^2 \end{aligned}$$

$$\operatorname{Var}(X) = E\left(x^2\right)-(E(x))^2$$

Proof

4. If *X* and *Y* are ***independent*** random variables, then prove Var (*X* +*Y*) = Var (*X*) + Var (*Y*) :

$$\operatorname{Var}(X) = E\left(x^2\right) - (E(x))^2$$

$$\begin{aligned}
\operatorname{Var}(X+Y) &= \mathrm{E}\left([X+Y]^2\right) - [\mathrm{E}(X+Y)]^2 \\
&\boxed{= \mathrm{E}\left(X^2 + 2XY + Y^2\right) - [\mathrm{E}(X) + \mathrm{E}(Y)]^2} \\
&= \mathrm{E}\left(X^2\right) + 2\,\mathrm{E}(XY) + \mathrm{E}\left(Y^2\right) - \mathrm{E}(X)^2 - 2\,\mathrm{E}(X)\,\mathrm{E}(Y) - \mathrm{E}(Y)^2 \\
&= \mathrm{E}\left(X^2\right) + 2\,\mathrm{E}(X)\,\mathrm{E}(Y) + \mathrm{E}\left(Y^2\right) - \mathrm{E}(X)^2 - 2\,\mathrm{E}(X)\,\mathrm{E}(Y) - \mathrm{E}(Y)^2 \\
&= \mathrm{E}\left(X^2\right) - \mathrm{E}(X)^2 + \mathrm{E}\left(Y^2\right) - \mathrm{E}(Y)^2 \\
&= \operatorname{Var}(X) + \operatorname{Var}(Y)
\end{aligned}$$

Expectation and Application

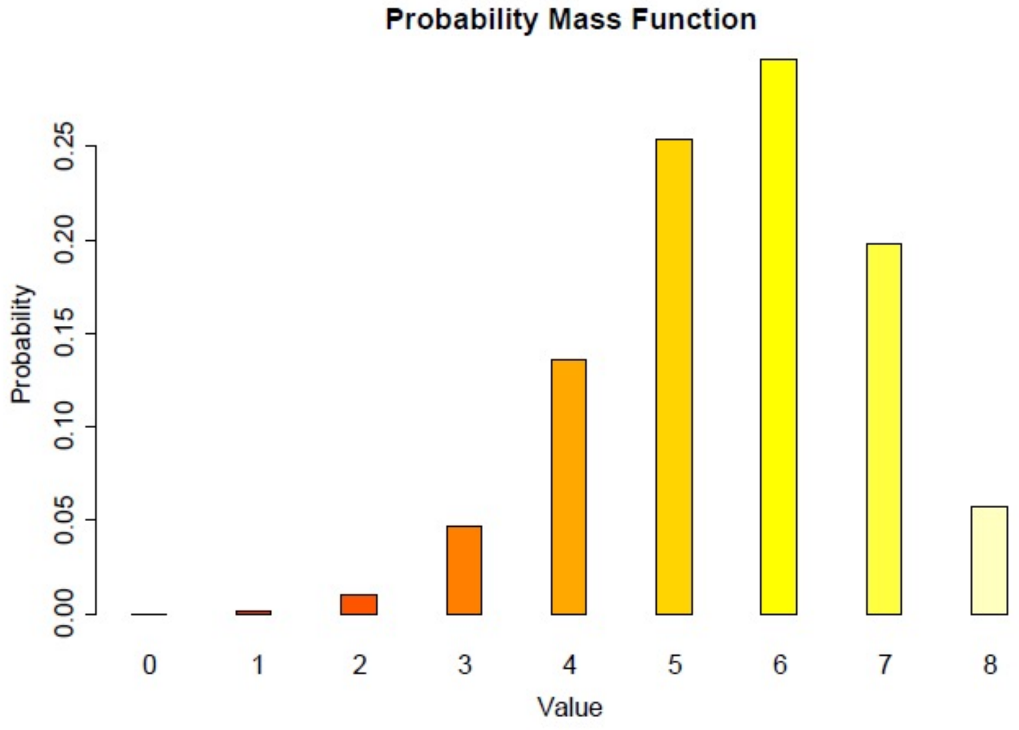
A certain kind of lizard lays 8 eggs, each of which will hatch independently with probability 0.7. Let Y denote the number of eggs which hatch. Then $Y \sim B(8, 0.7)$ (binomial distribution) . What is the PMF, CDF, expectation and variance of Y?

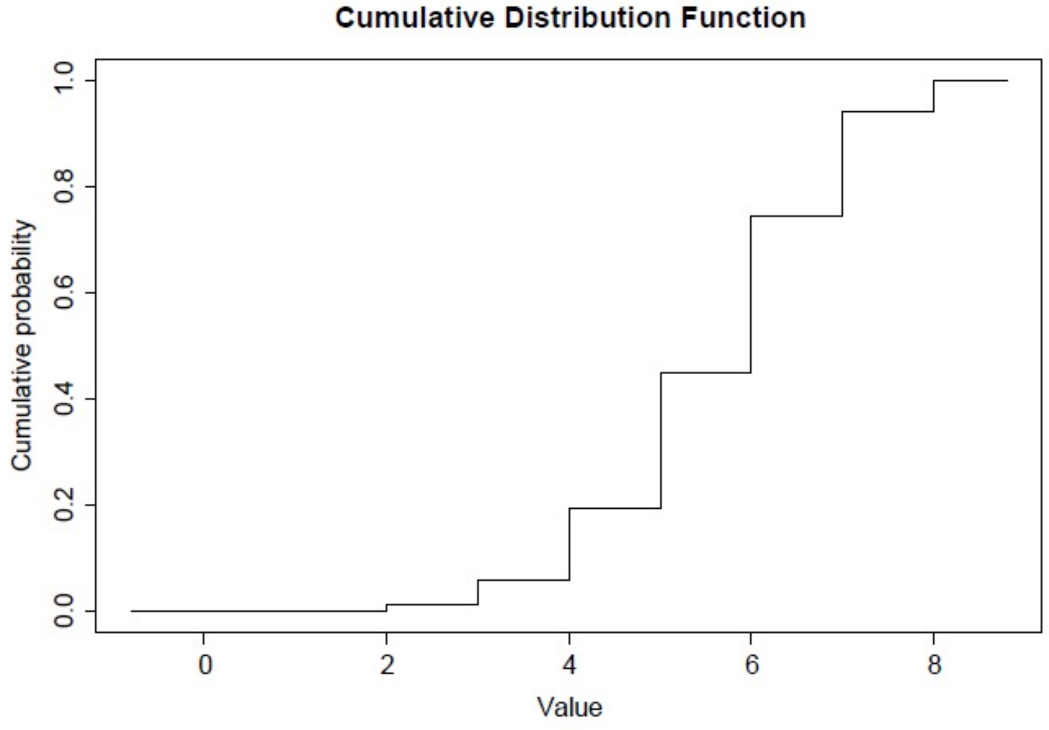
Solution: Since $Y \sim B(8, 0.7)$,

$$P(Y = k) = \binom{8}{k} 0.7^k 0.3^{8-k}, \; k = 0, 1, 2, \ldots, 8.$$

k	0	1	2	3	4	5	6	7	8
$P(Y = k)$	0.00	0.00	0.01	0.05	0.14	0.25	0.30	0.20	0.06
$F_Y(k)$	0.00	0.00	0.01	0.06	0.19	0.45	0.74	0.94	1.00

Expectation and Application





Expectation and Application

Expectation and variance:

Let I_j be the j-th egg which will hatch, and the I_j are mutually independent:

$$X = \sum_{j=1}^{n} I_j$$

$$\begin{aligned} \mathrm{E}(X) &= \mathrm{E}\left(\sum_{j=1}^{n} I_j\right) \\ &= \sum_{j=1}^{n} \mathrm{E}(I_j) \\ &= \sum_{j=1}^{n} p \\ &= np \end{aligned}$$

E(X) = 5.6

$$\begin{aligned} \mathrm{Var}(X) &= \mathrm{Var}\left(\sum_{j=1}^{n} I_j\right) \\ &= \sum_{j=1}^{n} \mathrm{Var}(I_j) \\ &= \sum_{j=1}^{n} p(1-p) \\ &= np(1-p). \end{aligned}$$

Var(X) = 1.68

Normal Distribution

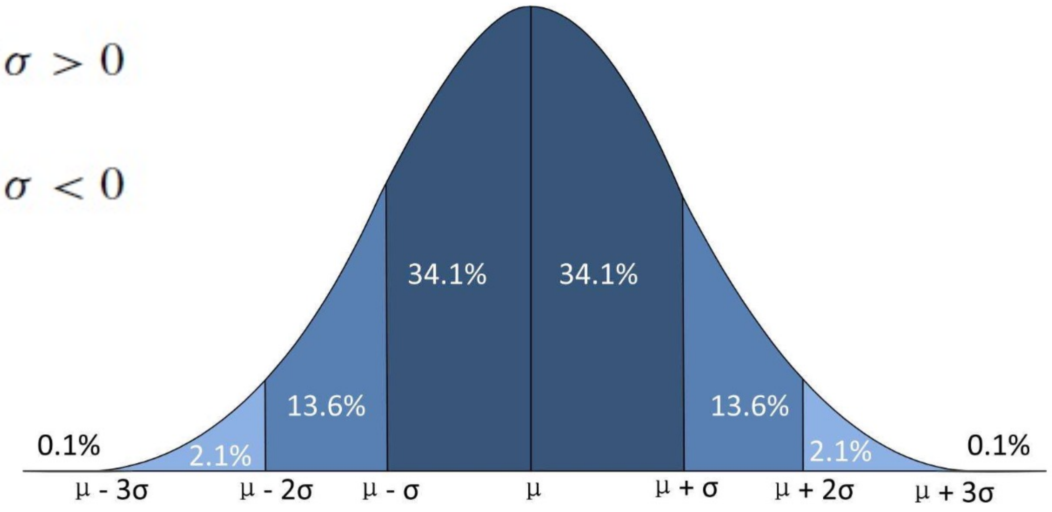
Let X have the standard normal distribution with density

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

Find the density of $Y = \sigma X + \mu$ for given constants μ and $\sigma \neq 0$. Also, find the density of $Z = X^2$.

$$\mathbf{P}(\sigma X + \mu \le y) = \mathbf{P}(\sigma X \le y - \mu) = \begin{cases} \mathbf{P}\left(X \le \dfrac{y-\mu}{\sigma}\right) & \text{if } \sigma > 0 \\ \mathbf{P}\left(X \ge \dfrac{y-\mu}{\sigma}\right) & \text{if } \sigma < 0 \end{cases}$$

$$= \begin{cases} F_X\left(\dfrac{y-\mu}{\sigma}\right) & \text{if } \sigma > 0 \\ 1 - F_X\left(\dfrac{y-\mu}{\sigma}\right) & \text{if } \sigma < 0 \end{cases}$$



Normal Distribution

differentiating with respect to y,

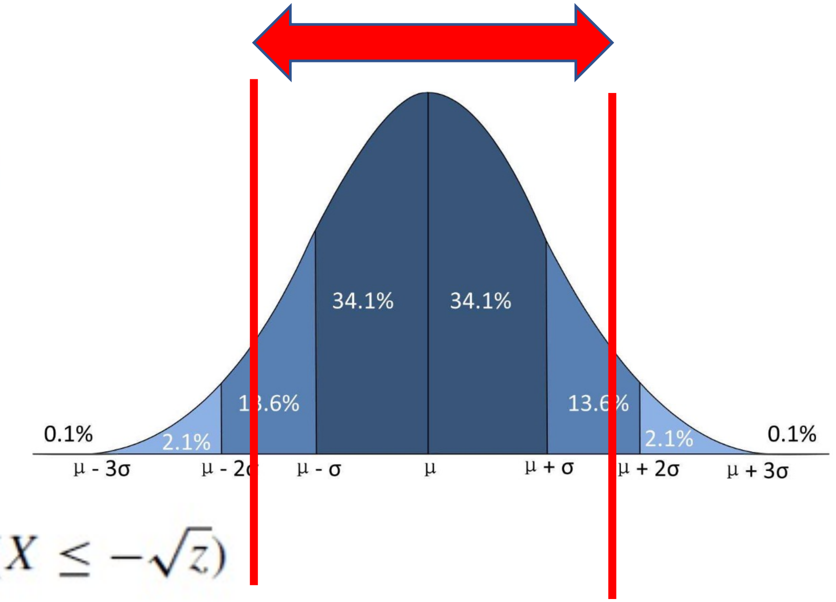
$$f_Y(y) = \frac{1}{|\sigma|} f_X\left(\frac{y-\mu}{\sigma}\right) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right).$$

f(X) ~ N(0,1) -> f(Y) : N(μ, σ)

$$\mathbf{P}(X^2 \le z) = \mathbf{P}(X \le \sqrt{z}) - \mathbf{P}(X \le -\sqrt{z}) = F_X(\sqrt{z}) - F_X(-\sqrt{z}).$$

$$f_Z(z) = \frac{1}{2\sqrt{z}} f_X(\sqrt{z}) + \frac{1}{2\sqrt{z}} f_X(-\sqrt{z}) = \frac{1}{\sqrt{2\pi z}} \exp\left(-\frac{1}{2}z\right).$$

f(X) ~ N(0,1) -> f(Z) : $\chi^2(1)$

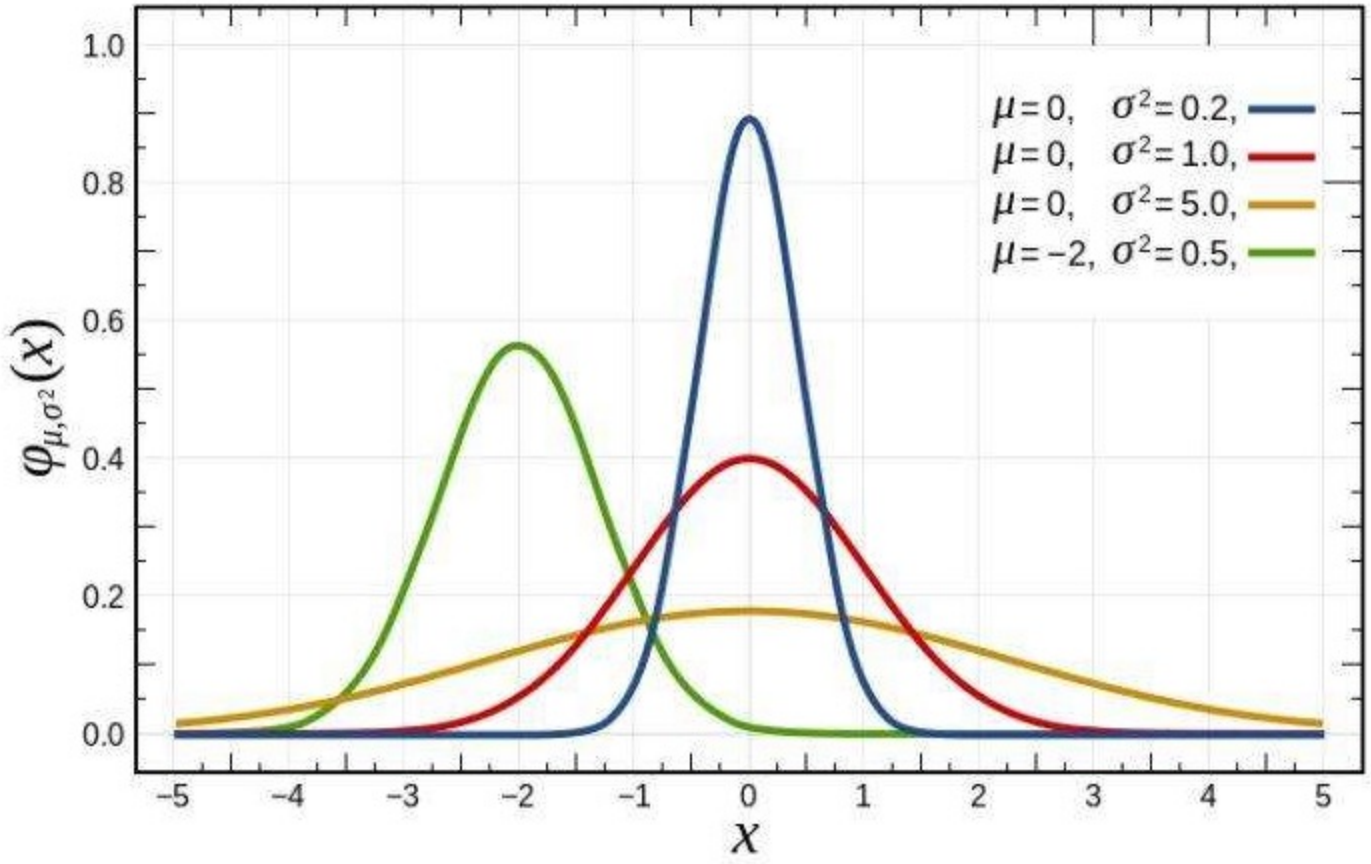


$\mathbf{P}(X \le -\sqrt{z})$

$\mathbf{P}(X \le \sqrt{z})$

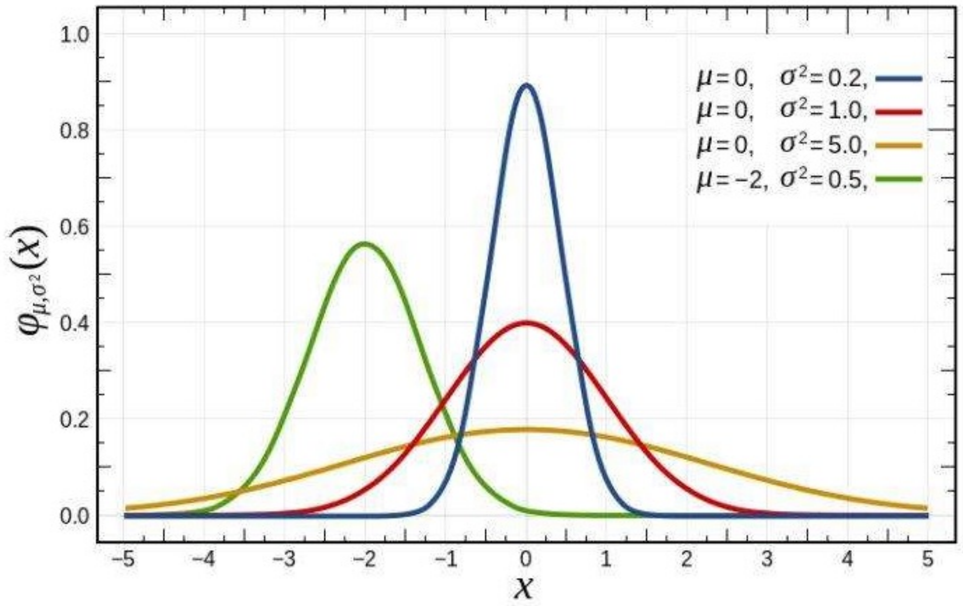
Normal Distribution

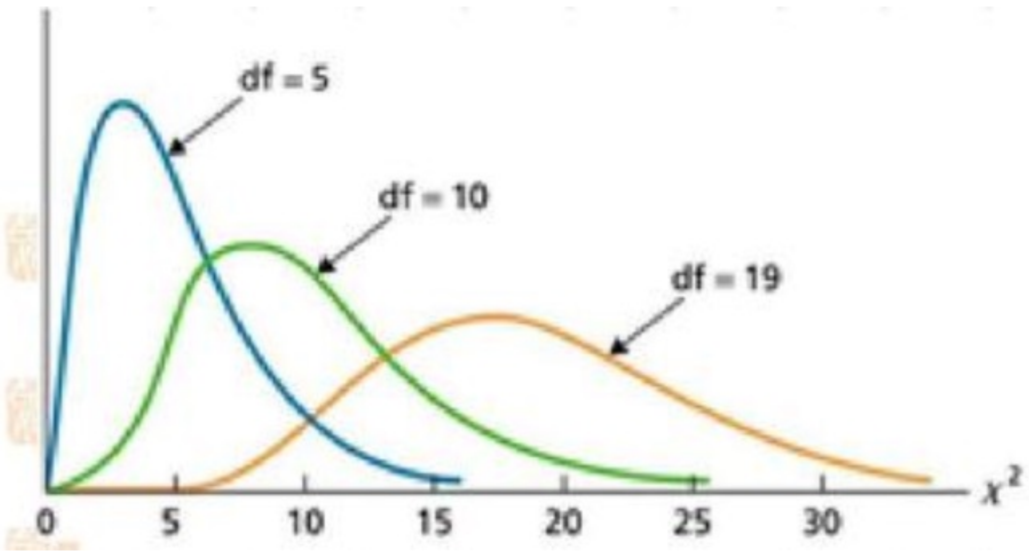
$f(X) \sim N(0, 1) \to f(Y) : N(\mu, \sigma)$



Normal Distribution

$f(X) \sim N(0,1) \rightarrow f(Z) : \chi^2(1)$





Remarks:

1. Prove/deduce the equations by yourselves.

2. Remember typical distributions and their properties.

3. Refer to Wikipedia and other online materials.

4. Use an appropriate distribution to model a question.

More

1. Why do we need to learn distributions?
   They help to model the realistic problem, e.g. toll a coin.

2. Why do we need statistics?
   They can provide general characteristics about a distribution/dataset, e.g. E(X): average capability, Var(X): extent of fluctuation.

3. What if we engage with a new/unknown distribution?
   Associate it with known distribution(s) so that we can directly make use the properties of statistics.