SAMSUNG

Together for Tomorrow!
Enabling People
Education for Future Generations

Samsung Innovation Campus

Artificial Intelligence Course

Together for Tomorrow! Enabling People
Education for Future Generations

Chapter 4.

Probability and Statistics

AI Course

Together for Tomorrow! Enabling People
Education for Future Generations

Chapter 4.

Probability and Statistics

Random Variable (1/3)

- Random variable: a function that assigns a number to the outcome of a random experiment.

  Ex) In a random experiment of coin flipping, assign 1 to the head (H) and assign 0 to the tail (H).

- There are two types of random variables:

  1) Discrete: for random experiments with a finite number of the possible outcomes.

     Ex) Random variable that represents coin flipping experiment. Possible outcomes $= \{H, T\}$.

     Ex) Random variable that represents dice rolling experiment. Possible outcomes $= \{1,2,3,4,5,6\}$.

  2) Continuous: for random experiments with an infinite number of the possible outcomes.

     Ex) Random variable that represents the heights of people.

     Ex) Random variable that represents the wages of people.

Random Variable (2/3)

Discrete probability distribution function $P(x)$:

- Maps the values of a discrete random variable to the corresponding probabilities.
- You will denote in upper case a random variable and in lower case a particular value of it.
  Ex) Given a random variable X,
  the probability of X taking on a value x is denoted by $P(X=x)$ or $P(x)$.

Continuous probability density function $f(x)$:

- When integrated, gives the interval probabilities $\Leftarrow$ More about this in the next Unit.

Random Variable (3/3)

- Properties of discrete probability distribution:

1) $0 \leq P(x) \leq 1$

2) $\sum_{all\ x_i} P(x_i) = 1$

Population and Sample (1/2)

Population:

- The entirety of the data set subject to the analysis.
- Can be either "real" or "idealized".
- The properties of a population are called parameters:

  Ex) Mean, standard deviation, variance, etc. of a population.

Sample:

- A subset of the population.
- The properties of a sample are called statistics:

  Ex) Mean, standard deviation, variance, etc. of a sample.

Population and Sample (2/2)

- We could calculate the mean, standard deviation, variance, etc. using a probability distribution function.
  - These quantities are properties of an “idealized” group.
  - They can be interpreted as parameters of an idealized population.

Population Mean (1/2)

- Population mean is denoted by μ.
  - It is also called the expected value of a random variable: $E[X]$.

    1) For the discrete case with $P(x)$= probability distribution function:

$$\mu = E[X] = \sum_{all\ x} x\,P(x)$$

    2) For the continues case with $f(x)$= probability density function:

$$\mu = E[X] = \int x\,f(x)dx$$

Population Mean (2/2)

Properties of the population mean:

1) $E[c] = c$

2) $E[c\,X] = c\,E[X]$

3) $E[X + c] = E[X] + c$

4) $E[X + Y] = E[X] + E[Y]$

Note: c stands for a constant.

Population Variance (1/2)

- The population variance is denoted by σ^2 or $Var(X)$.

1) For the discrete case with $P(x)$= probability distribution function:

$$\sigma^2 = Var(X) = \sum_{all\ x} (x-\mu)^2\, P(x)$$

2) For the continues case with $f(x)$= probability density function:

$$\sigma^2 = Var(X) = \int (x-\mu)^2 f(x)dx$$

Note: In both cases $\sigma^2 = E[X^2] - (E[X])^2$

Note: The population standard deviation $\sigma = \sqrt{\sigma^2}$

Together for Tomorrow! Enabling People
Education for Future Generations

Population Variance (2/2)

Properties of the population variance:

1) $Var(c) \ = 0$

2) $Var(X + c) = Var(X)$

3) $Var(c\, X) = c^{2}\, Var(X)$

4) $Var(X + Y) = Var(X) + Var(Y) + 2\, Cov(X, Y)$

$\qquad = Var(X) + Var(Y)$ $\Leftarrow$ Only when X and Y are independent to each other!

Note: c stands for a constant.

Chapter 4.

Probability and Statistics

Together for Tomorrow! Enabling People

Education for Future Generations

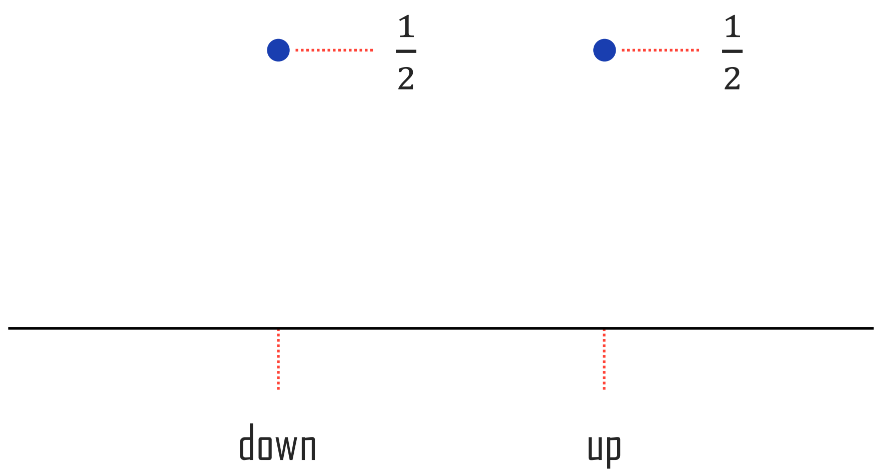
Discrete Probability Distribution (1/19)

- Bernoulli random variable and probability distribution:

Discrete Probability Distribution (2/19)

Bernoulli random variable and probability distribution:

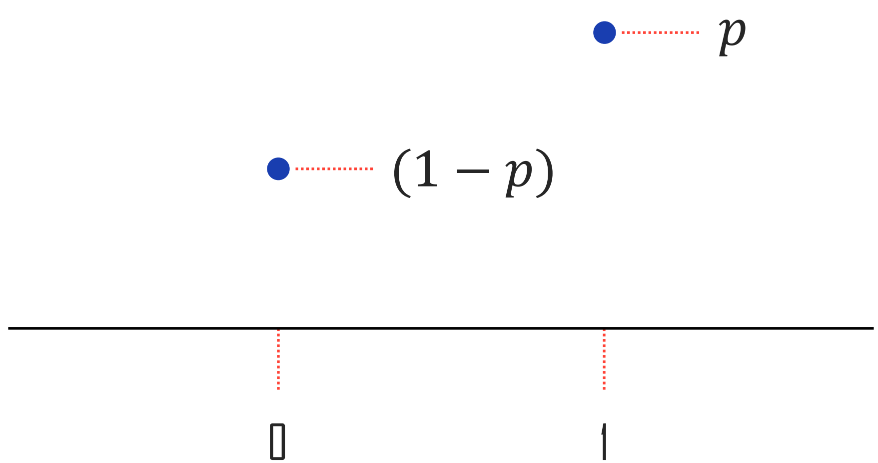
Ex) Coin flipping:



Discrete Probability Distribution (3/19)

Bernoulli random variable and probability distribution:

Ex) General binary outcome situation:



Discrete Probability Distribution (4/19)

Bernoulli random variable and probability distribution:

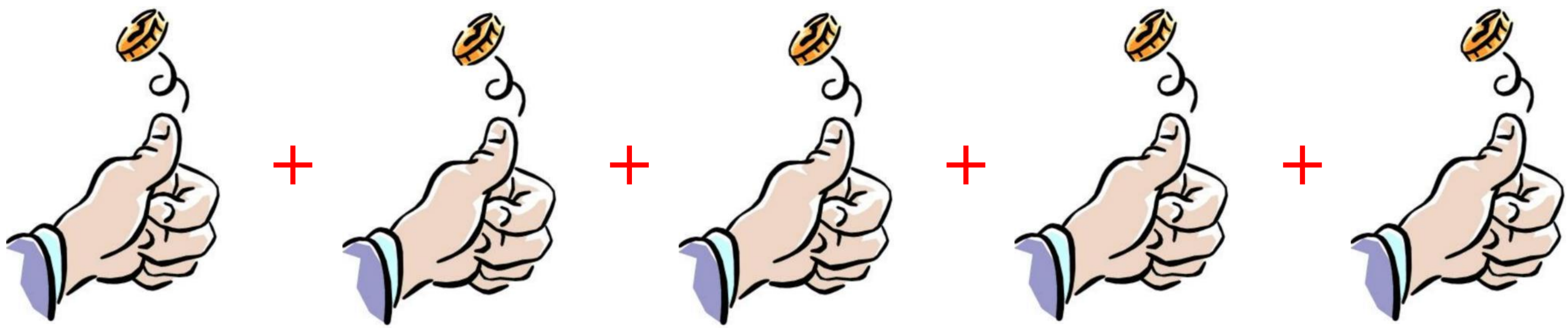
1) "X is a Bernoulli random variable" $\Leftrightarrow$ $X \sim Ber(p)$

2) $P(x) = p^x (1-p)^{1-x}$

- Mean = p
- Variance = $p\,(1-p)$
- Standard deviation = $\sqrt{p\,(1-p)}$

Discrete Probability Distribution (5/19)

- Binomial random variable and probability distribution:



Discrete Probability Distribution (6/19)

Binomial random variable and probability distribution:

1) "X is a Binomial random variable" $\Leftrightarrow X \sim Bin(n, p)$

2) $X_{bin} = \underbrace{X_{Ber} + X_{Ber} + \cdots + X_{Ber}}_{\leftarrow \quad n \quad \rightarrow}$

3) $P(x) = \binom{n}{x} p^x (1-p)^{n-x}$ $\Leftarrow 0 \leq x \leq n$

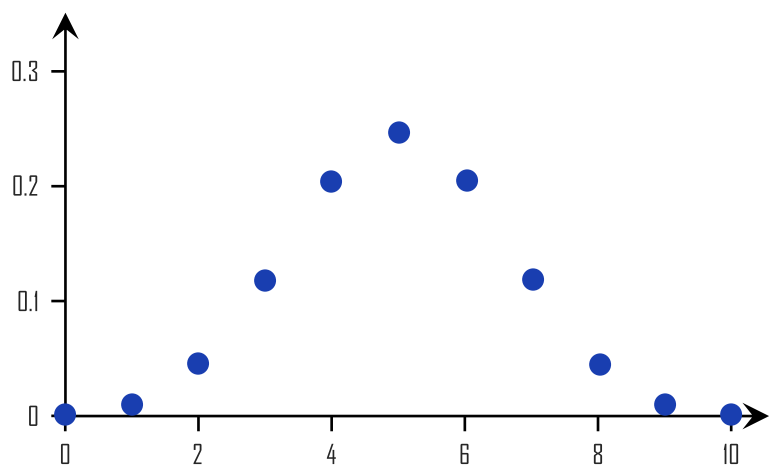
- Mean $= n\,p$
- Variance $= n\,p\,(1-p)$
- Standard deviation $= \sqrt{n\,p\,(1-p)}$

Together for Tomorrow! Enabling People
Education for Future Generations

Discrete Probability Distribution (7/19)

Binomial random variable and probability distribution:

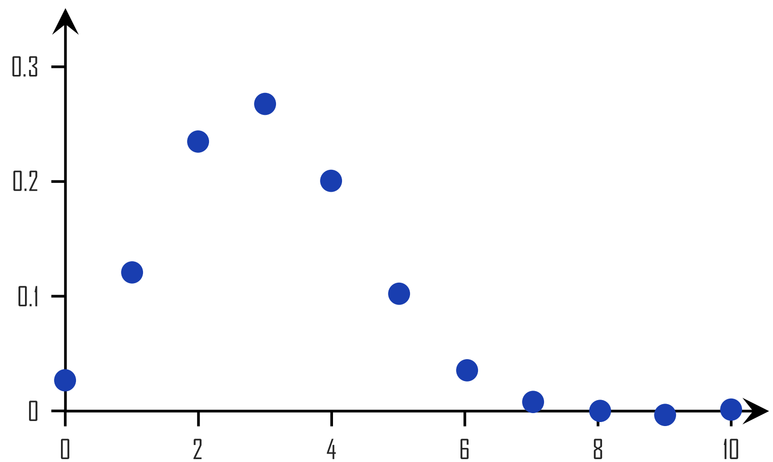
$$n = 10, p = 0.5$$



Discrete Probability Distribution (8/19)

- Binomial random variable and probability distribution:

$$n = 10, p = 0.3$$

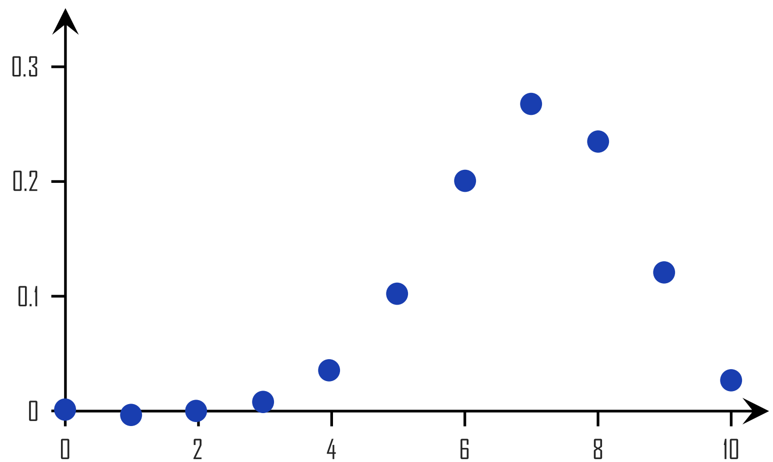


Together for Tomorrow! Enabling People
Education for Future Generations

Discrete Probability Distribution (9/19)

- Binomial random variable and probability distribution:

$$n = 10, p = 0.7$$



Discrete Probability Distribution (10/19)

Binomial random variable and probability distribution:

Ex) It is known that the daily probability of raining is 20% for the following five days.

What is the probability of zero rainy day?

$$P(0) = \binom{5}{0} 0.2^0 (1-0.2)^{5-0} = \frac{5!}{0!\ 5!} 0.8^5 = 0.8^5 = 0.328$$

Discrete Probability Distribution (11/19)

- Binomial random variable and probability distribution:

Ex) It is known that the daily probability of raining is 20% for the following five days.

What is the probability of exactly two rainy days?

$$P(2) = \binom{5}{2} 0.2^2 (1-0.2)^{5-2} = \frac{5!}{2!\,3!} 0.2^2 \times 0.8^3 = 10 \times 0.2^2 \times 0.8^3 = 0.205$$

Discrete Probability Distribution (12/19)

Binomial random variable and probability distribution:

Ex) It is known that the daily probability of raining is 20% for the following five days.

What is the probability of two or fewer rainy days?

$$P(X \leq 2) = P(0) + P(1) + P(2)$$

$$= \binom{5}{0} 0.2^0 (1-0.2)^{5-0} + \binom{5}{1} 0.2^1 (1-0.2)^{5-1} + \binom{5}{2} 0.2^2 (1-0.2)^{5-2}$$

$$= 1 \times 1 \times 0.8^5 + 5 \times 0.2 \times 0.8^4 + 10 \times 0.2^2 \times 0.8^3$$

$$= 0.328 + 0.41 + 0.205$$

$$= 0.942$$

Discrete Probability Distribution (13/19)

Poisson random variable and probability distribution:

- Named after the French scientist Simeon D. Poisson.
- Describes the frequency of events for a given interval of time or space.

  Ex) Number of emails received per hour.

  Ex) Number of earthquakes per year,

  Ex) Number of chocolate chips in a cookie.

Discrete Probability Distribution (14/19)

Poisson random variable and probability distribution:

1) "X is a Poisson random variable" ⇔ $X \sim Pois(\lambda)$

2) $P(x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad \Leftarrow \; 0 \leq x$

- Mean $=\lambda$
- Variance $=\lambda$
- Standard deviation $= \sqrt{\lambda}$

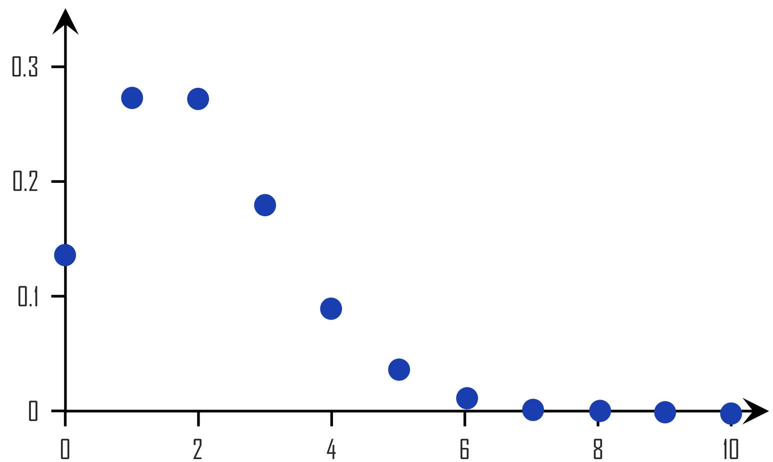
3) By increasing n while decreasing p the Binomial distribution converges to the Poisson with $\lambda = n\,p$.

Together for Tomorrow! Enabling People
Education for Future Generations

Discrete Probability Distribution (15/19)

- Poisson random variable and probability distribution:

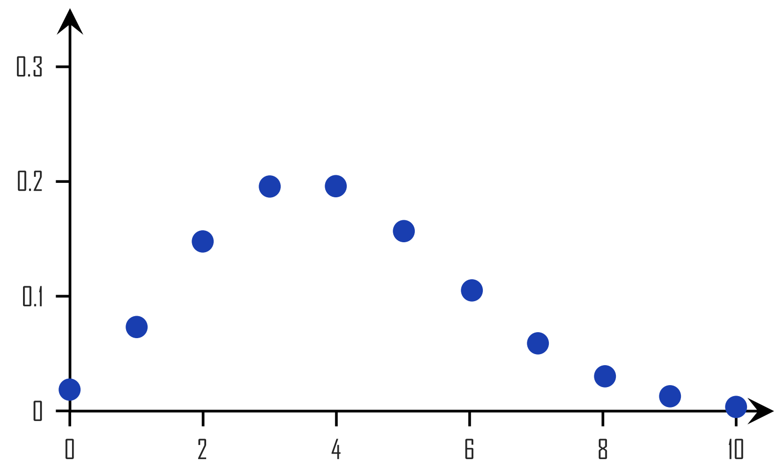
$$lambda = 2$$



Discrete Probability Distribution (16/19)

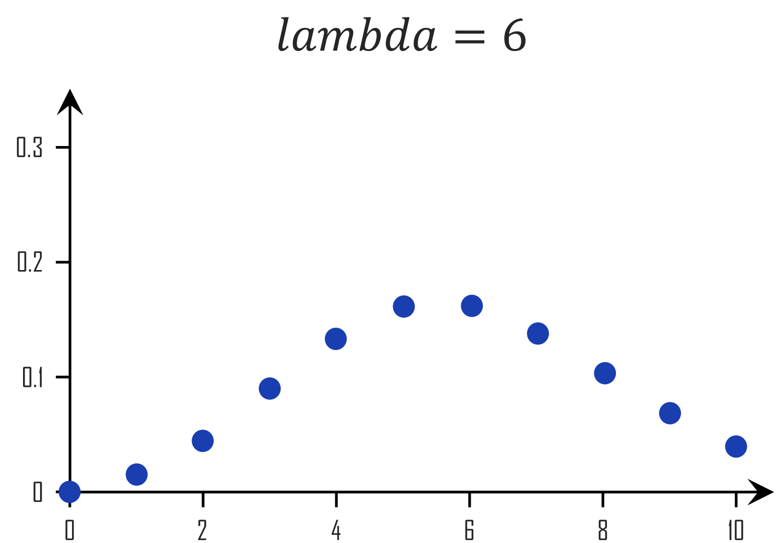
- Poisson random variable and probability distribution:

$lambda = 4$



Discrete Probability Distribution (17/19)

- Poisson random variable and probability distribution:



Discrete Probability Distribution (18/19)

- Poisson random variable and probability distribution:

Ex) During the last 100 days 40 spam mails were received.

What is the probability that tomorrow there will be zero spam mail?

You can calculate that $\lambda = \frac{40}{100} = 0.4$

Then, $P(0) = \frac{0.4^0 e^{-0.4}}{0!} = e^{-0.4} = 0.6703$

Discrete Probability Distribution (19/19)

- Poisson random variable and probability distribution:

Ex) During the last 100 days 40 spam mails were received.

What is the probability that tomorrow there will be more than one spam mails?

You can calculate that $\lambda = \frac{40}{100} = 0.4$

Then, do $P(1) + P(2) + P(3) + \cdots$?? $\Rightarrow$ Hard!

Instead, you can calculate by doing $1 - P(0) = 1 - 0.6703 = 0.3297$

Together for Tomorrow! Enabling People
Education for Future Generations

Coding Exercise #0301

Follow practice steps on 'ex_0301.ipynb' file.

Together for Tomorrow! Enabling People

Education for Future Generations

Chapter 4.

Probability and Statistics

UNIT 2.

Understanding of Statistics I

Samsung Innovation Campus

Unit 2.
Understanding of Statistics I

What this unit is about:

- You will learn about the continuous probability densities.
- You will learn about the conjoint probability.
- You will learn about the correlation and linear relationship.

Expected outcome:

- Ability to define and calculate the probability.
- Ability to calculate the expected values.
- Ability to model random events with the appropriate probability density.

How to check your progress:

- Coding Exercises.
- Quiz.

Chapter 4.

Probability and Statistics

Together for Tomorrow! Enabling People

Education for Future Generations

Continuous Probability Density (1/24)

Continuous random variable and probability density:

- Infinite number of possible values.
- The probability at a specific value is zero: $P(X = \boldsymbol{x_0}) = 0$
- Non-zero probability only for intervals: $P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x)dx$
- Cumulative probability: $CDF(x) = P(-\infty < X \leq x)$

$$= \int_{-\infty}^{x} f(y)dy$$

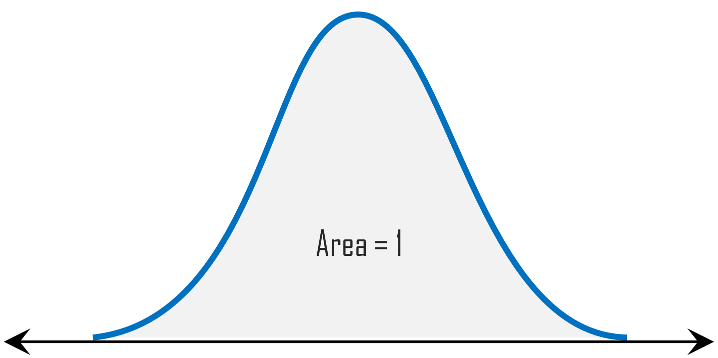
- We can calculate $P(x_1 \leq X \leq x_2) = CDF(x_2) - CDF(x_1)$

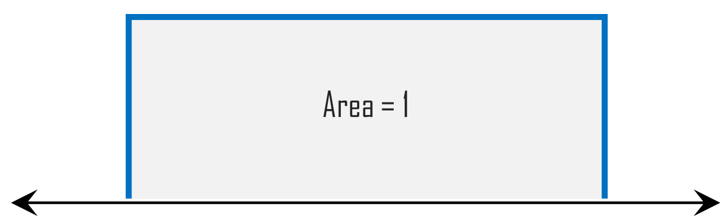
Continuous Probability Density (2/24)

Properties of continuous probability density:

1) $0 \leq f(x)$

2) $\int f(x)\, dx = 1$





Continuous Probability Density (3/24)

Uniform random variable and probability density:

1) "X is a Uniform random variable in the interval $[a, b]$" $\Leftrightarrow X \sim Unif(a, b)$

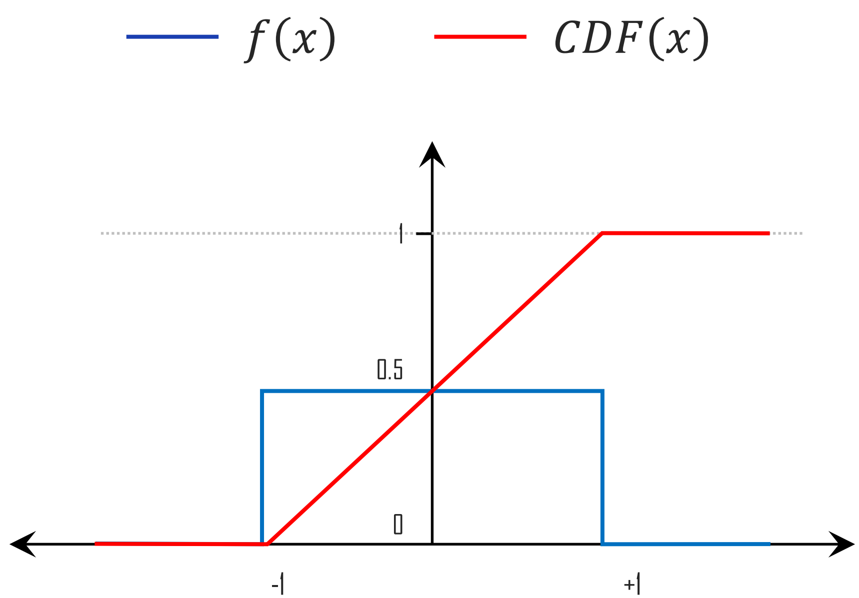
2) $f(x) = \frac{1}{(b-a)}$ $\Leftarrow$ defined in the interval $[a, b]$ and zero elsewhere.

- Mean= $\frac{1}{2}(a+b)$
- Variance= $\frac{1}{12}(b-a)^2$
- Standard deviation= $\frac{1}{\sqrt{12}}(b-a)$

Continuous Probability Density (4/24)

Uniform random variable and probability density:

Ex) $a = -1, b = +1$



Continuous Probability Density (5/24)

- Uniform random variable and probability density:

Ex) The lifespan of a light bulb is uniformly distributed between 10,000 and 20,000 hours.

What is the probability that the light bulb will last between 12,000 and 15,000 hours of usage?

$$P(12000 \leq X \leq 15000) = \int_{12000}^{15000} \frac{1}{20000 - 10000} dx = 0.3$$

Together for Tomorrow! Enabling People
Education for Future Generations

Continuous Probability Density (6/24)

Uniform random variable and probability density:

Ex) The lifespan of a light bulb is uniformly distributed between 10,000 and 20,000 hours.

What is the probability that the light bulb will last 15,000 hours or more?

$$P(15000 \leq X \leq 20000) = \int_{15000}^{20000} \frac{1}{20000 - 10000} dx = 0.5$$

Continuous Probability Density (7/24)

Normal random variable and probability density:

1) "X is a Normal random variable with mean μ and variance σ^2" $\Leftrightarrow X \sim N(\mu, \sigma^2)$

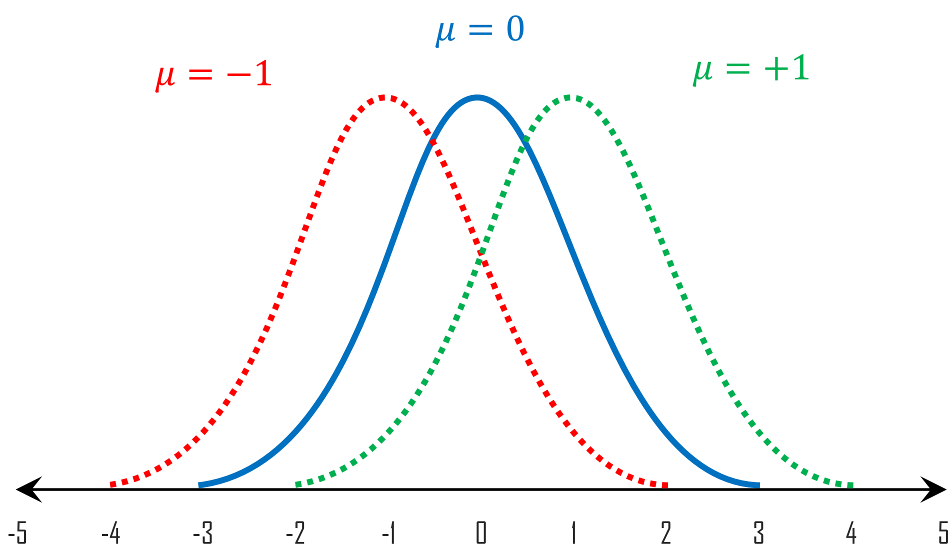
2) $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $\Leftarrow$ defined in the interval $(-\infty, +\infty)$

- Mean $= \mu$
- Variance $= \sigma^2$
- Standard deviation $= \sigma$

Together for Tomorrow! Enabling People
Education for Future Generations

Continuous Probability Density (8/24)

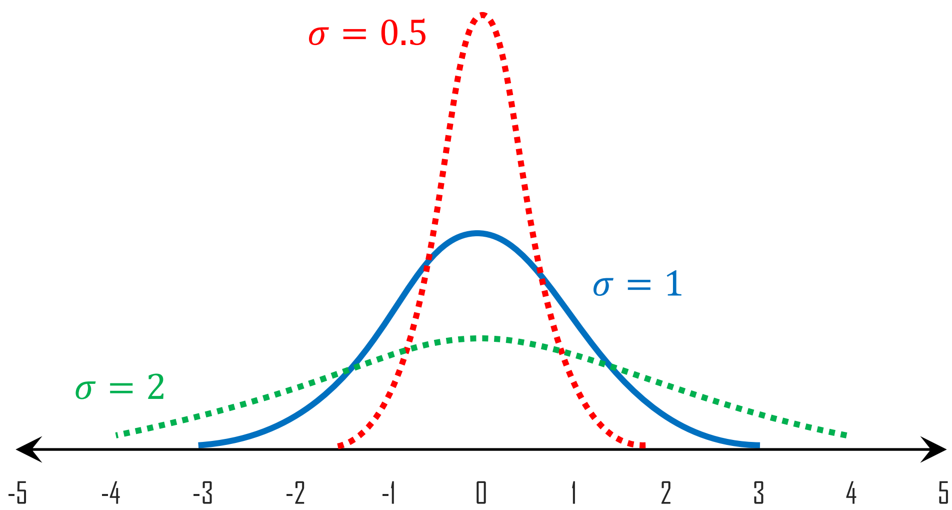
- Normal random variable and probability density:



- μ is the so-called "location" parameter.

Continuous Probability Density (9/24)

Normal random variable and probability density:



- σ is the so-called "shape" parameter.

Continuous Probability Density (10/24)

Standard Normal random variable and probability density:

1) "X is a Standard Normal random variable" $\Leftrightarrow X \sim N(0,1)$

2) $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ $\Leftarrow$ defined in the interval $(-\infty, +\infty)$

- Mean = 0
- Variance = 1
- Standard deviation = 1

Together for Tomorrow! Enabling People
Education for Future Generations

Continuous Probability Density (11/24)

Normal random variable and probability density: "Addition and Subtraction"

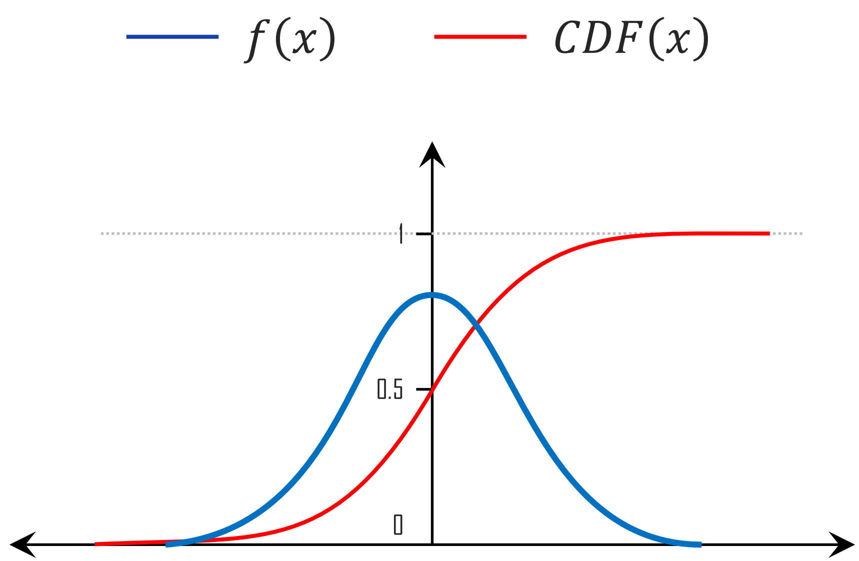
- Suppose that $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$.
- And, also suppose that X and Y are independent to each other.

1) $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

2) $X - Y \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$

Continuous Probability Density (12/24)

- Standard Normal random variable and probability density:



Continuous Probability Density (13/24)

- Suppose that we have a random variable $X \sim N(\mu, \sigma^2)$, then it is possible to convert to another random variable Z such that $Z \sim N(0,1)$. This conversion process is called "standardization".

$$Z = \frac{X - \mu}{\sigma}$$

- Standardized values are called "z-scores":

$$\text{z-score} = \frac{x - \mu}{\sigma}$$

- It is possible to go in the opposite direction: from the Standard Normal to the Normal.

$$X = \sigma Z + \mu$$

Continuous Probability Density (14/24)

Standard Normal random variable and probability density:

Ex) A barista takes in average 50 seconds to make a cup of coffee with the standard deviation of 20 seconds. What is the probability that the barista will take 48 ~ 54 seconds for your next cup of coffee? Use the table of the Standard Normal CDFs.

1) First of all, let us standardize:

$$x_1 = 48 \Rightarrow z_1 = \frac{x_1-\mu}{\sigma} = \frac{48-50}{20} = -\frac{2}{20} = -0.1$$

$$x_2 = 54 \Rightarrow z_2 = \frac{x_2-\mu}{\sigma} = \frac{54-50}{20} = \frac{4}{20} = 0.2$$

z	CDF(z)
-0.2	0.4207
-0.1	0.4602
0	0.5
0.1	0.5398
0.2	0.5793

Continuous Probability Density (15/24)

Standard Normal random variable and probability density:

Ex) A barista takes in average 50 seconds to make a cup of coffee with the standard deviation of 20 seconds. What is the probability that the barista will take 48 ~ 54 seconds for your next cup of coffee? Use the table of the Standard Normal CDFs.

2) $P(z_1 \leq Z \leq z_2) = CDF(z_2) - CDF(z_1)$

$= CDF(0.2) - CDF(-0.1)$

$= 0.5793 - 0.4602 = 0.1191$

z	CDF(z)
-0.2	0.4207
-0.1	0.4602
0	**0.5**
0.1	0.5398
0.2	0.5793

Together for Tomorrow! Enabling People
Education for Future Generations

Continuous Probability Density (16/24)

Chi-square random variable and probability density:

1) The Chi-square random variable Q is the sum of squares of $\boldsymbol{k}$ independent Standard Normal

random variables X_i: $Q = X_1^2 + X_2^2 + \cdots + X_k^2$

2) "X is a Chi-square random variable with the degree of freedom $\boldsymbol{k}$" $\Leftrightarrow$ $Q \sim \chi^2(k)$

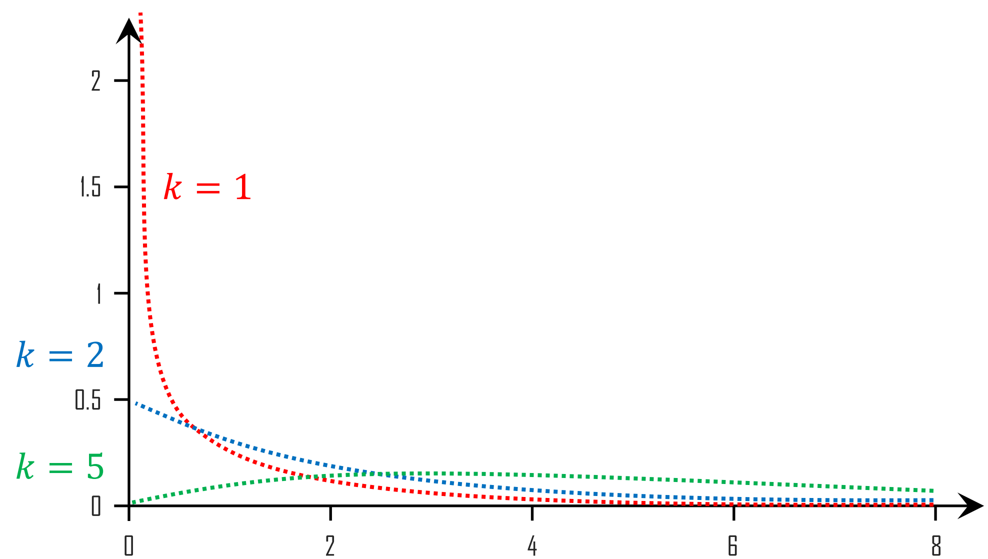
3) $f(x) = \frac{1}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})} x^{\frac{k}{2}-1}$ $\Leftarrow$ defined in the interval $(0, +\infty)$

- Mean $= k$
- Variance $= 2k$
- Standard deviation $= \sqrt{2k}$

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x}\, dx \qquad z > 0.$$

Continuous Probability Density (17/24)

- Chi-square random variable and probability density:



Continuous Probability Density (18/24)

- Chi-square random variable and probability density: "Addition"

Suppose that $Q_1 \sim \chi^2(\boldsymbol{k_1})$ and $Q_2 \sim \chi^2(\boldsymbol{k_2})$ which are independent to each other.

Then, $Q_1 + Q_2 \sim \chi^2(\boldsymbol{k_1} + \boldsymbol{k_2})$.

- The proof is straightforward:

$$Q_1 + Q_2 = \underbrace{\{X^2 + X^2 + \cdots + X^2\}}_{\leftarrow\ \boldsymbol{k_1}\ \rightarrow} + \underbrace{\{X^2 + \cdots + X^2\}}_{\leftarrow\ \boldsymbol{k_2}\ \rightarrow}$$

Thus, by regrouping $Q_1 + Q_2 \sim \chi^2(\boldsymbol{k_1} + \boldsymbol{k_2})$

Together for Tomorrow! Enabling People
Education for Future Generations

Continuous Probability Density (19/24)

Student-t random variable and probability density:

1) When $Q \sim \chi^2(\nu)$ and $Z \sim N(0,1)$, then the Student-t random variable is defined as following:

$$T = \frac{Z}{\sqrt{Q/\nu}}$$

2) "T is a Student-t random variable with the degree of freedom ν" ⇔ $T \sim t(\nu)$

3) $f(x) = \dfrac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\dfrac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$ ⇐ defined in the interval $(-\infty, +\infty)$

- Mean = 0 Variance = $\dfrac{\nu}{\nu-2}$ Standard deviation = $\sqrt{\dfrac{\nu}{\nu-2}}$ ⇐ when $\nu > 2$

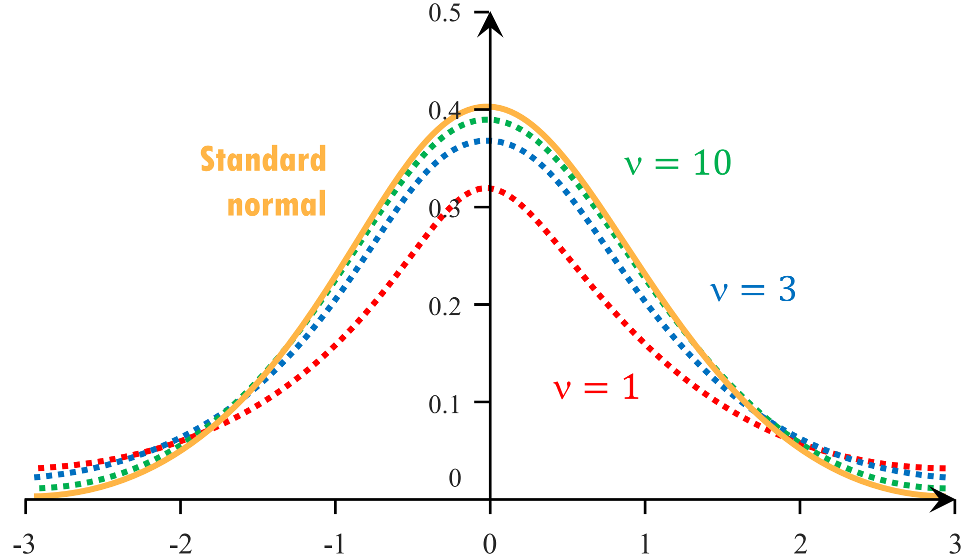
Continuous Probability Density (20/24)

- Student-t random variable and probability density:
  - Useful when doing the interval estimation with small sample size.
  - Useful also when we do the hypothesis testing of the means.
  - The degree of freedom and the sample size are related: $\nu = n - 1$
  - As the degree of freedom (~sample size) increases, the Student-t converges to the Standard Normal.

Together for Tomorrow! Enabling People
Education for Future Generations

Continuous Probability Density (21/24)

- Student-t random variable and probability density:



- As the degree of freedom ν increases, the Student-t converges to the standard normal.

Continuous Probability Density (22/24)

F random variable and probability density:

1) When $Q_1 \sim \chi^2(d_1)$ and $Q_2 \sim \chi^2(d_2)$, then the F random variable is defined as following:

$$X = \frac{Q_1/d_1}{Q_2/d_2}$$

2) "X is a F random variable with the degrees of freedom d_1 and d_2" $\Leftrightarrow X \sim F(d_1, d_2)$
3) d_1 = degree of freedom for the numerator and d_2 = the degree of freedom for the denominator.

- Mean $= \frac{d_2}{d_2 - 2}$ $\Leftarrow$ when $d_2 > 2$
- Variance $= \frac{2d_2^2(d_1 + d_2 - 2)}{d_1(d_2 - 4)(d_2 - 2)^2}$ $\Leftarrow$ when $d_2 > 4$
- Standard deviation $= \sqrt{\frac{2d_2^2(d_1 + d_2 - 2)}{d_1(d_2 - 4)(d_2 - 2)^2}}$ $\Leftarrow$ when $d_2 > 4$

Continuous Probability Density (23/24)

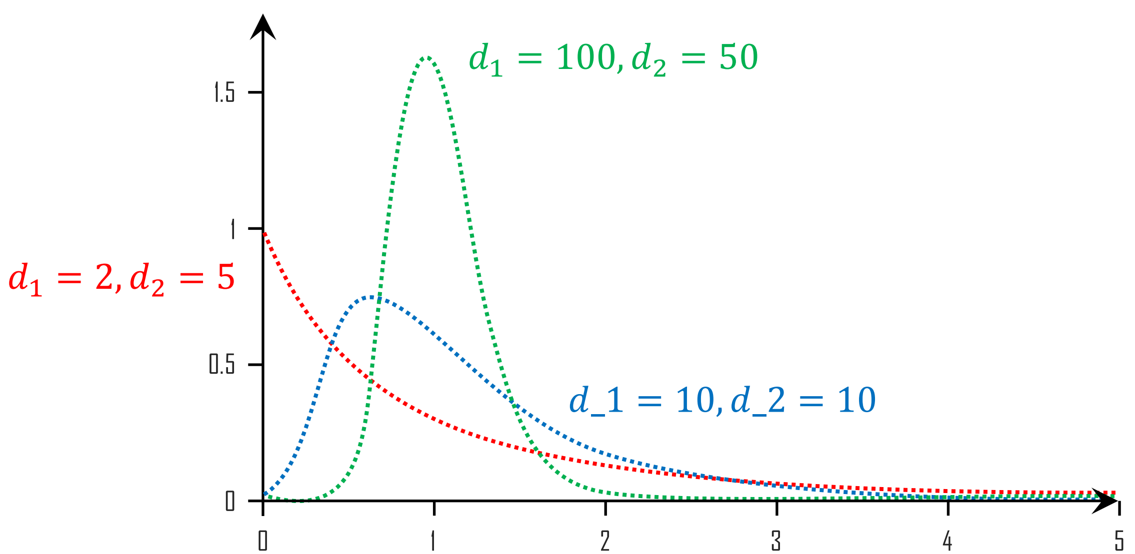
F random variable and probability density:

- Useful when comparing variances of two samples.
- Also used for ANOVA (Analysis of Variance) that compares the group means.
- As d_2 increases, the F random variable converges to ${}^{Q_1}/_{d_1}$.

Together for Tomorrow! Enabling People
Education for Future Generations

Continuous Probability Density (24/24)

- F random variable and probability density:



- There are two degrees of freedom: d_1 and d_2.

Coding Exercise #0302

Follow practice steps on 'ex_0302.ipynb' file.

Together for Tomorrow! Enabling People
Education for Future Generations

Chapter 4.

Probability and Statistics

Conjoint Probability (1/13)

- Bivariate conjoint probability: for two random variables X and Y.
- For the discrete case: $P(x,y) = P(X=x, Y=y)$
  - The probability when $X = x$ **AND** $Y = y$
- For the continuous case: $f(x,y) = f(X=x, Y=y)$
  - The probability density when $X = x$ **AND** $Y = y$

Conjoint Probability (2/13)

Covariance:

- About two variables X and Y with an implicit conjoint probability.
- Can be calculated as expected value: $Cov(X,Y)=E[(X-\mu\_x)(Y-\mu\_y)]$
- A rather "simpler" expression: $Cov(X,Y)=E[X\ Y]-E[X]\ E[Y]$
- When $X = Y$, we have $Cov(X,X)=Var(X)$

Conjoint Probability (3/13)

Correlation:

- Correlation and covariance are related: $Cor(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$
- Value in the interval $[-1,1]$.
- Correlation is a measure of the linear relationship between X and Y.

a) $Cor(X,Y) > 0$: a positive linear relationship between X and Y.

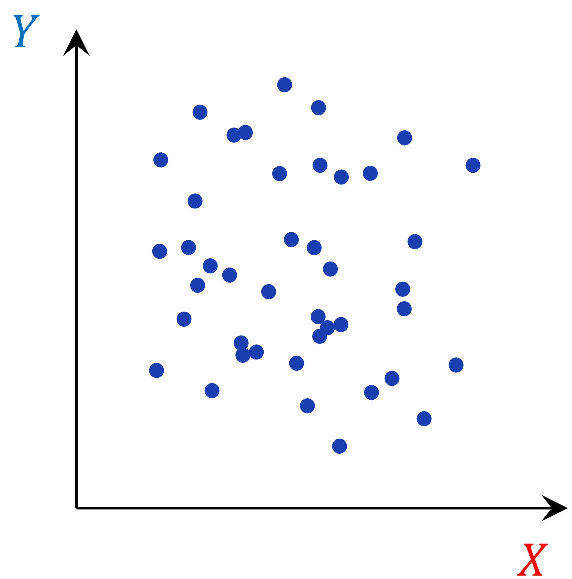
b) $Cor(X,Y) < 0$: a negative linear relationship between X and Y.

c) $Cor(X,Y) = 0$: no linear relationship between X and Y.

Together for Tomorrow! Enabling People
Education for Future Generations

Conjoint Probability (4/13)

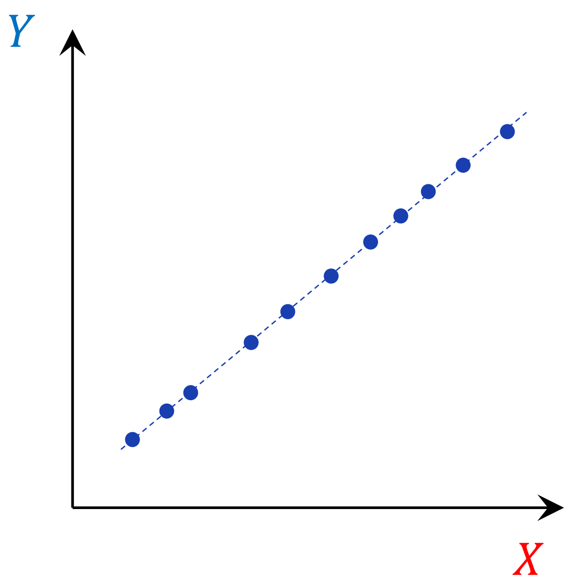
- Correlation:



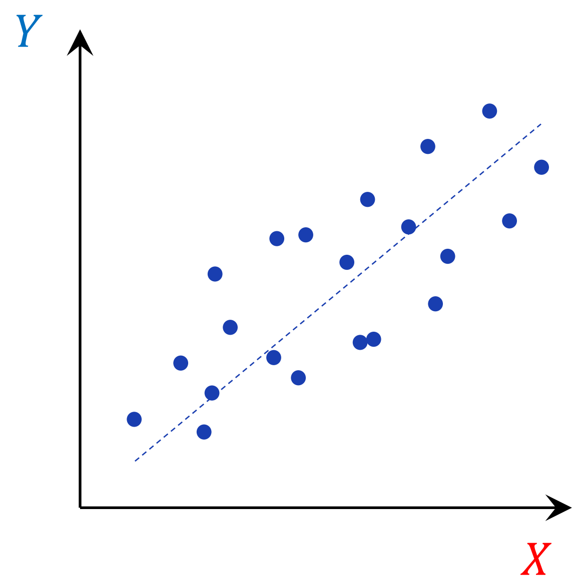
No linear relationship

Conjoint Probability (5/13)

- Correlation:



Strong positive correlation

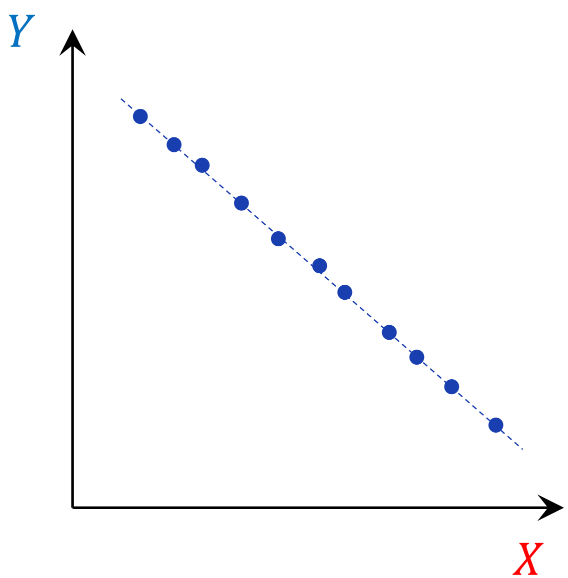


Weak positive correlation

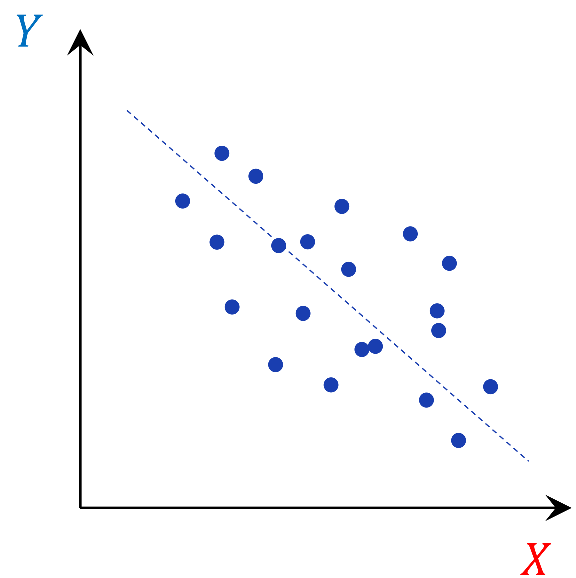
Together for Tomorrow! Enabling People
Education for Future Generations

Conjoint Probability (6/13)

Correlation:



Strong negative correlation

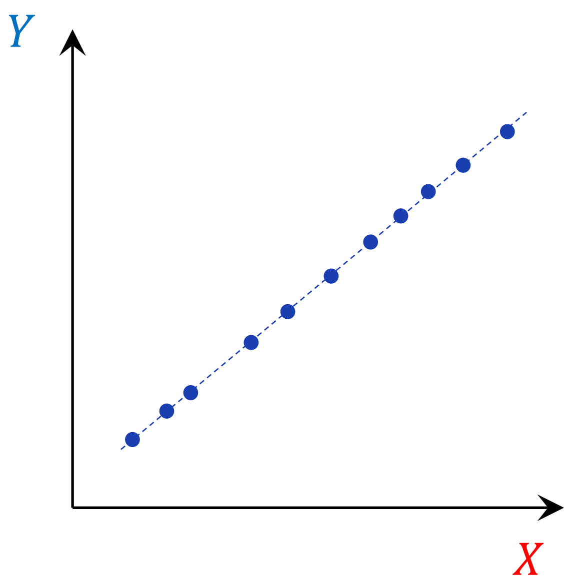


Weak negative correlation

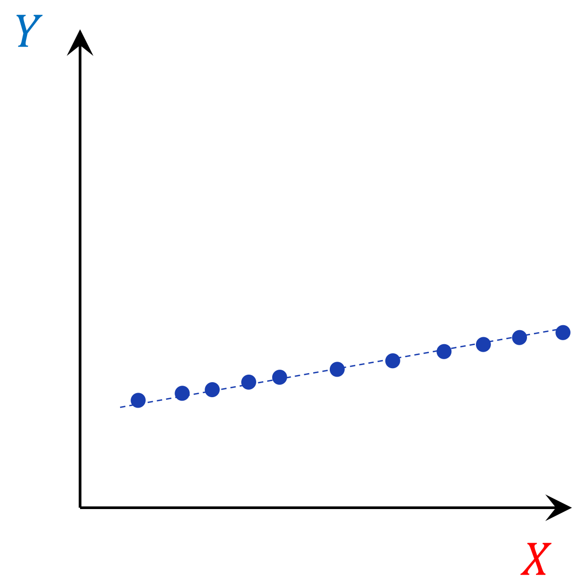
Together for Tomorrow! Enabling People
Education for Future Generations

Conjoint Probability (7/13)

- Correlation:



Strong positive correlation

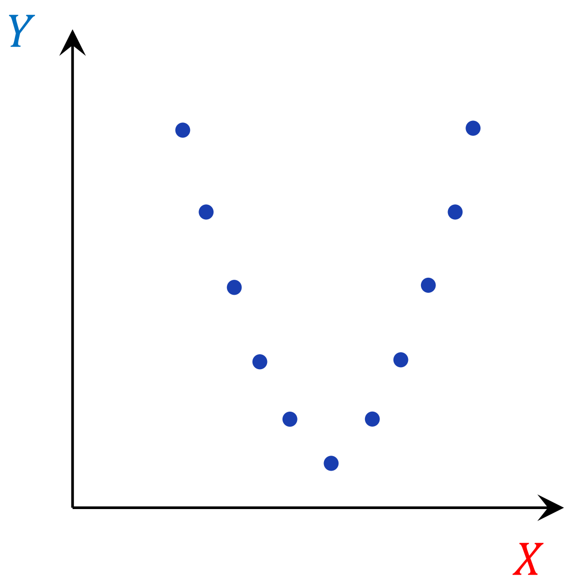


Strong positive correlation

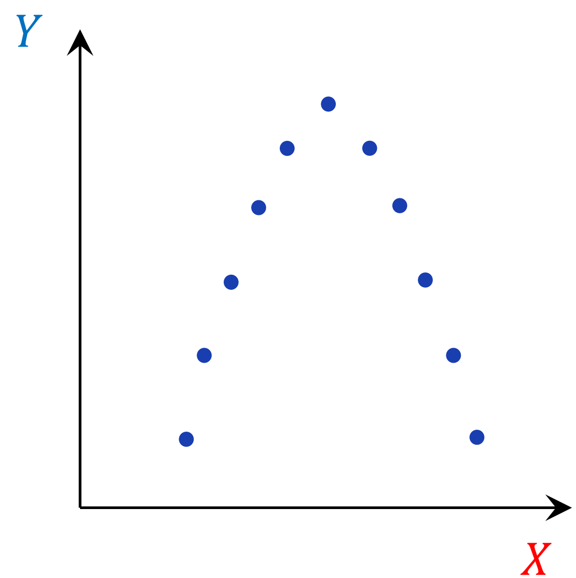
Together for Tomorrow! Enabling People
Education for Future Generations

Conjoint Probability (8/13)

- Correlation:



No linear relationship



No linear relationship

Conjoint Probability (9/13)

Dependence and Correlation:

- If X and Y are independent: $P(X,Y)=P(X)\ P(Y)$ or $f(X,Y)=f(X)\ f(Y)$

  Then, $Cov(X,Y)=E[X\ Y]-E[X]E[Y]=E[X]E[Y]-E[X]E[Y]=0$.

  As $Cor(X,Y)= \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$ we conclude that $Cor(X,Y)=0$.

  So, independence **always** means uncorrelatedness.
- However, X and Y being uncorrelated (*) cannot guarantee the independence all the time.

(*) that is $Cor(X,Y)=0$

Conjoint Probability (10/13)

Dependence and Correlation:

Ex) Let's consider tossing two coins represented by the random variables X and Y.

The sample space is composed of HH, HT, TH, TT.

a) The individual probability distributions are:

$$\left.\begin{array}{l} P(X=H)=\frac{1}{2} \\ P(X=T)=\frac{1}{2} \end{array}\right\} P(X)$$

$$\left.\begin{array}{l} P(Y=H)=\frac{1}{2} \\ P(Y=T)=\frac{1}{2} \end{array}\right\} P(Y)$$

Together for Tomorrow! Enabling People
Education for Future Generations

Conjoint Probability (11/13)

Dependence and Correlation:

Ex) Let's consider tossing two coins represented by the random variables X and Y.

The sample space is composed of HH, HT, TH, TT.

b) In the conjoint probability, you can check that indeed X and Y are independent:

$$P(X{=}H, Y{=}H) = \frac{1}{4} = P(X{=}H) \times P(Y{=}H)$$

$$P(X{=}H, Y{=}T) = \frac{1}{4} = P(X{=}H) \times P(Y{=}T)$$

$$P(X{=}T, Y{=}H) = \frac{1}{4} = P(X{=}T) \times P(Y{=}H)$$

$$P(X{=}T, Y{=}T) = \frac{1}{4} = P(X{=}T) \times P(Y{=}T)$$

Together for Tomorrow! Enabling People
Education for Future Generations

Conjoint Probability (12/13)

Dependence and Correlation:

Ex) The random variable X has equal probability 1/3 for the each one of the outcomes -1, 0, 1.

And, there is another random variable Y defined as $Y=X^2$.

Then, the conjoint probabilities are as following:

$$P(X=-1, Y=1) = \frac{1}{3}, \quad P(X=0, Y=0) = \frac{1}{3}, \quad P(X=1, Y=1) = \frac{1}{3}$$

1) Let us calculate the $Cor(X,Y)$:

$$E[X] = -1 \times P(X=-1) + 0 \times P(X=0) + 1 \times P(X=1) = -\frac{1}{3} + 0 + \frac{1}{3} = 0$$

$$E[Y] = 0 \times P(Y=0) + 1 \times P(Y=1) = 0 + 1 \times \frac{2}{3} = \frac{2}{3}$$

$$E[X\,Y] = E[X\,X^2] = E[X^3] = -1 \times P(X=-1) + 0 \times P(X=0) + 1 \times P(X=1) = 0$$

So, $Cov(X,Y) = E[X\,Y] - E[X]\;E[Y] = 0 - 0 \times \frac{2}{3} = 0$ **and** $Cor(X,Y) = 0$ too.

Conjoint Probability (13/13)

Dependence and Correlation:

Ex) The random variable X has equal probability 1/3 for the each one of the outcomes -1, 0, 1.

And, there is another random variable Y defined as $Y=X^2$.

Then, the conjoint probabilities are as following:

$$P(X=-1, Y=1) = \frac{1}{3} \quad , \quad P(X=0, Y=0) = \frac{1}{3} \quad , \quad P(X=1, Y=1) = \frac{1}{3}$$

2) Now that you know X and Y are uncorrelated, check for the independence.

$$P(X=-1, Y=1) = \frac{1}{3} \neq P(X=-1) \times P(Y=1) = \frac{1}{3} \times \frac{2}{3} = \frac{2}{9}$$

$$P(X=0, Y=0) = \frac{1}{3} \neq P(X=0) \times P(Y=0) = \frac{1}{3} \times \frac{1}{3} = \frac{1}{9}$$

$$P(X=1, Y=1) = \frac{1}{3} \neq P(X=1) \times P(Y=1) = \frac{1}{3} \times \frac{2}{3} = \frac{2}{9}$$

Summarizing $P(X,Y) \neq P(X)P(Y)$ so that X and Y are **not** independent.

Here, we found a case where X and Y are uncorrelated but not independent!

Together for Tomorrow! Enabling People
Education for Future Generations

Chapter 4.

Probability and Statistics

UNIT 3.

Understanding of Statistics II

Unit 3. Understanding of Statistics II

What this unit is about:

- You will learn about descriptive statistics.
- You will learn about the central limit theorem.
- You will learn about point estimation and interval estimation.

Expected outcome:

- Ability to summarize the statistical properties of a sample.
- Ability to differentiate the population from the sample.
- Ability to apply interval estimation.

How to check your progress:

- Coding Exercises.
- Quiz.

Chapter 4.

Probability and Statistics

Together for Tomorrow! Enabling People
Education for Future Generations

Together for Tomorrow! Enabling People
Education for Future Generations

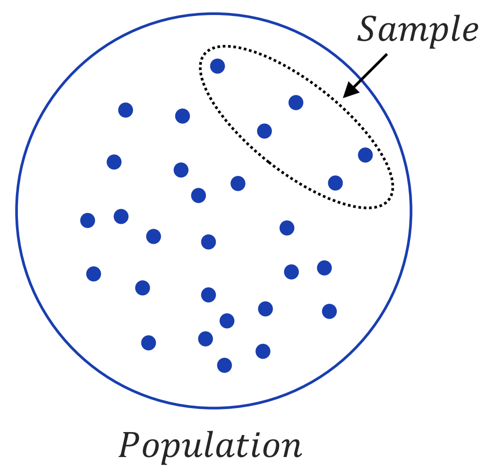
Population vs Sample

Population

- Whole data set subject to the analysis. It can be real or conceptual.

Sample:

- A subset of the population.
- You analyze samples because the population is hard or impossible to reach.
- You would like to draw conclusions about the population using its samples.



Descriptive vs Inferential

- Descriptive statistics:
  - Summarizes the data without generalization as a primary goal.
  - Extracts the properties of data as is: sample statistics.
- Inferential statistics:
  - Analysis of the samples with the purpose of making generalized statements about the population.
- The main difference between descriptive and inferential statistics lies in the goal.

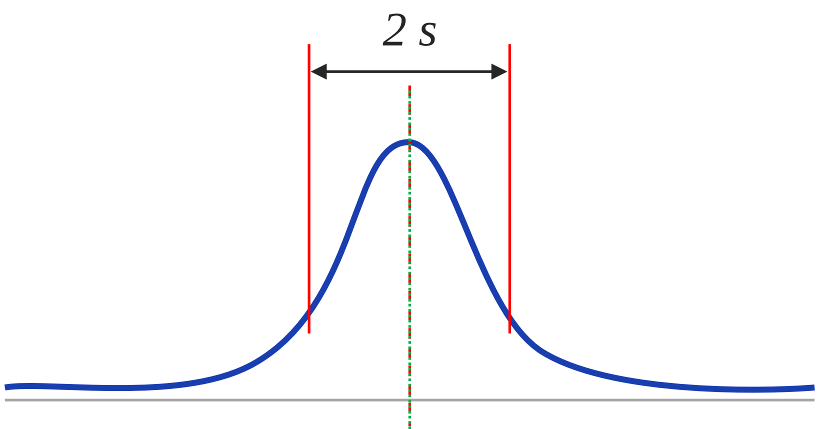
Descriptive Statistics (1/10)

Sample statistics:

- Mean value: $\overline{x}$
- Median: m
- Variance: s^2
- Standard deviation: $s = \sqrt{s^2}$
- Covariance: s_{XY}
- Correlation: r
- Skewness: ς
- Kurtosis: κ

Descriptive Statistics (2/10)

Sample statistics:



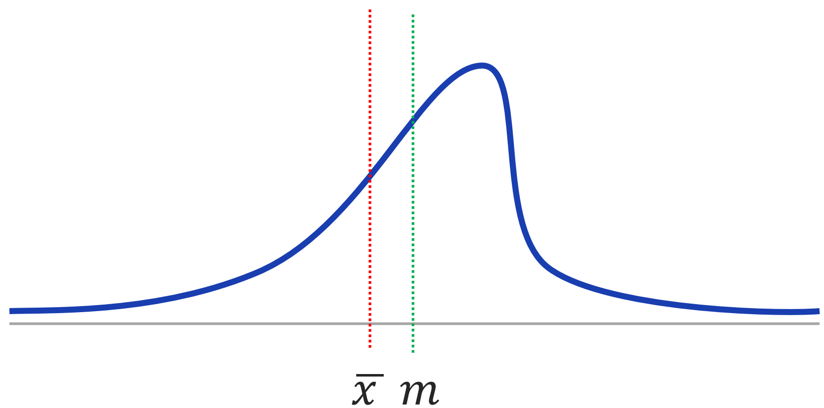
$$\bar{x} = m$$

$$\varsigma = 0$$

Symmetric

Descriptive Statistics (3/10)

- Sample statistics:



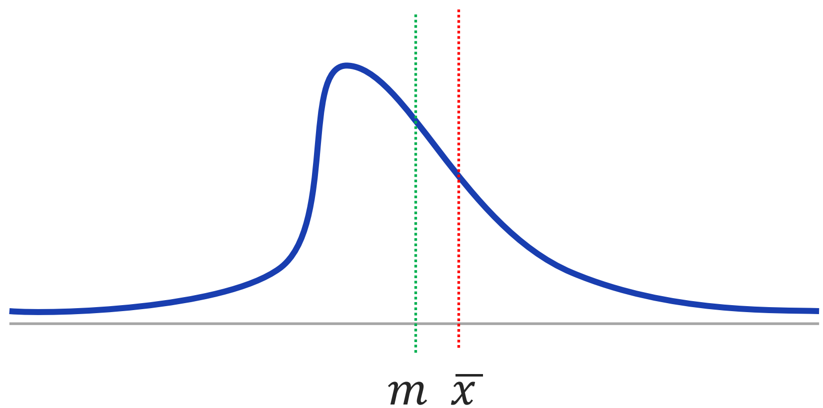
$\varsigma < 0$

Asymmetric

Together for Tomorrow! Enabling People
Education for Future Generations

Descriptive Statistics (4/10)

- Sample statistics:



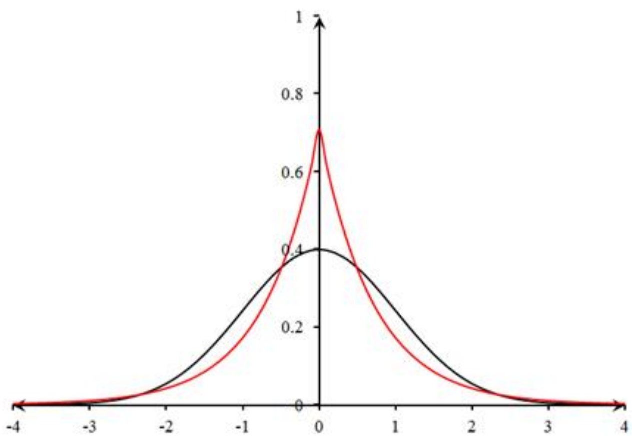
$$\varsigma > 0$$

Asymmetric

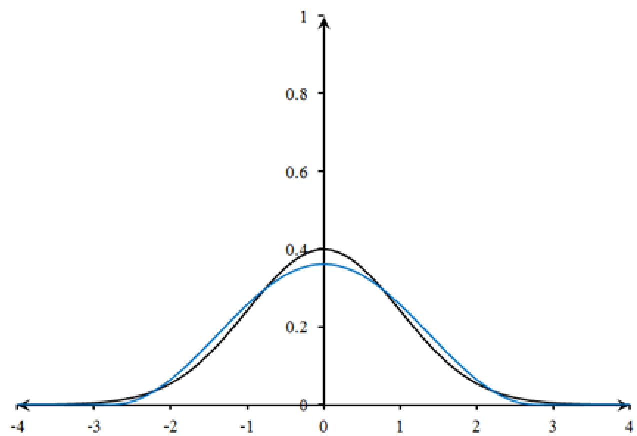
Together for Tomorrow! Enabling People
Education for Future Generations

Descriptive Statistics (5/10)

- Sample statistics:



$\kappa > 0$



$\kappa < 0$

Together for Tomorrow! Enabling People
Education for Future Generations

Descriptive Statistics (6/10)

Sample statistics:

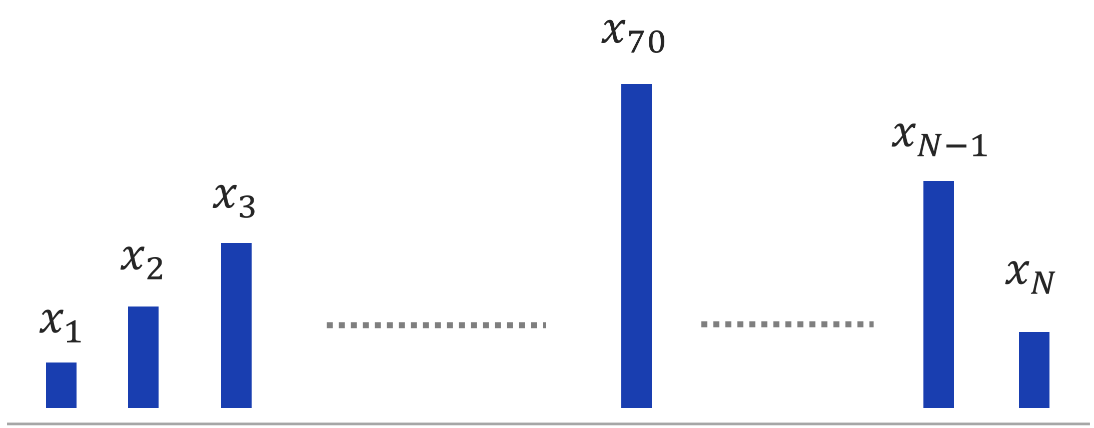
- Sample variance: $s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$
- Sample covariance: $s_{XY} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}$
- Sample correlation: $r = \frac{s_{XY}}{s_X s_Y}$

Together for Tomorrow! Enabling People
Education for Future Generations

Descriptive Statistics (7/10)

Sample statistics: quantile

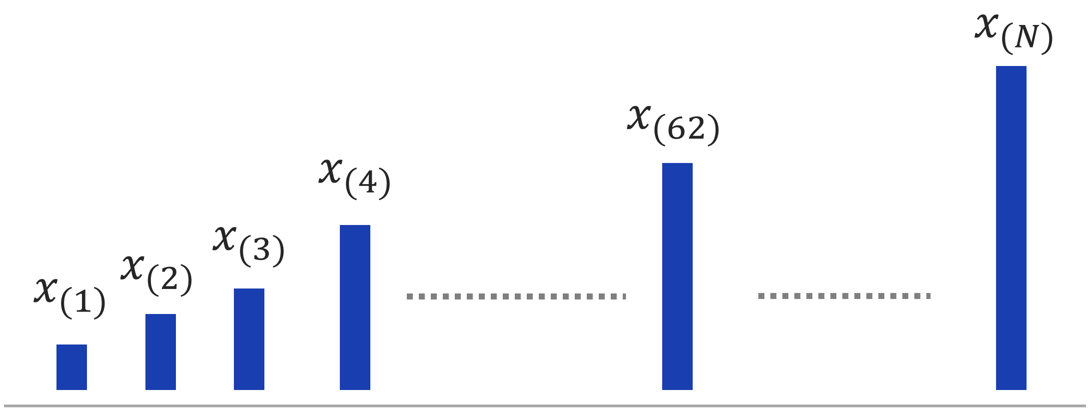
- Let's suppose that you have a sample of values $x_1, x_2, x_3, \cdots, x_N$. You can represent them as bars.



Descriptive Statistics (8/10)

Sample statistics: quantile

- Now, let us sort the values from the smallest to the largest and get $x_{(1)}, x_{(2)}, x_{(3)}, \cdots, x_{(N)}$.



- $\alpha\ quantile$= value at the position $int(N \times \alpha)$

Descriptive Statistics (9/10)

Sample statistics: quantile

- Quantile: α is probability 0 ~ 1.
- Percentile: α is given as percentage 0% ~ 100%.

  → Minimum = 0% percentile.

  → Maximum = 100% percentile.

  $$\text{Range} = \text{Maximum} - \text{Minimum}$$
- Quartile: α subdivided into four equal intervals.

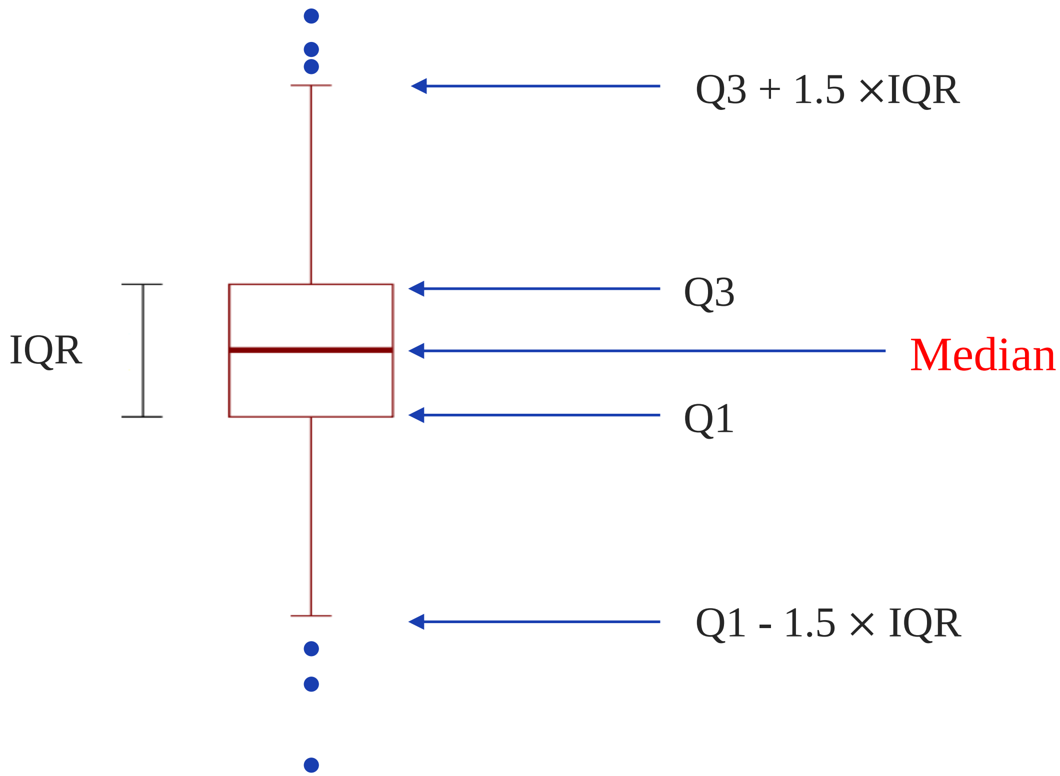
  → 1st quartile (Q1): 25% percentile.

  → 2nd quartile (Q2): 50% percentile = Median.

  → 3rd quartile (Q3): 75% percentile.

Descriptive Statistics (10/10)

- Boxplot:



Coding Exercise #0303

Follow practice steps on 'ex_0303.ipynb' file.