Probability Distribution

Dr Wan Nor Arifin
Unit of Biostatistics and Research Methodology, Universiti Sains Malaysia.
wnarifin@usm.my

Last update: 23 September, 2018

Wan Nor Arifin, 2018. Probability Distribution by Wan Nor Arifin is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-sa/4.0/.

Outlines

Introduction........1
Random variable........1
Discrete random variable........1
Continuous random variable........2
Probability Distributions of Discrete Random Variables........2
Probability Distribution........2
Cumulative Distribution Function (cdf)........2
Bernoulli trial........3
Binomial Distribution........4
Poisson Distribution........6
Probability Distributions of Continuous Random Variables........8
Probability Distribution........8
Cumulative Distribution Function (cdf)........9
Normal Distribution........9
Standard Normal Distribution........11
Topics for self-study........13
References........13
Online resources........14

Introduction

Random variable

- It is a function from sample space Ω to the real numbers.
- Random outcomes → random numbers.

Discrete random variable

- It is a random variable that can take only a finite or at most a countable infinite number of values (Rice, 1995).
- It is characterized by gaps or interruptions in the values that it can assume (Daniel, 1995).
- Real number, count, frequency.
- Disease grading normal, mild, moderate, severe → 0, 1, 2, 3 – finite.
- Number of car accident in August 2012 → 0, 1, 2, 3 … – infinite.
- 20 cent coin tossed *n* times, frequency of Hibiscus *minus* frequency of Tepak Sirih (can have negative number).

Continuous random variable

- It is a random variable that can take on a continuum of values (Rice, 1995).
- It does not possess the gaps or interruptions in the values that it can assume (Daniel, 1995).
- Decimal place.
- Weight → 0 kg to … kg – ratio scale, true zero.
- Temperature → *x* ºC to *y* ºC – interval scale, no true zero.

Probability Distributions of Discrete Random Variables

Probability Distribution

- A **probability mass function (pmf)** (also known as **frequency function**) – for discrete random variable.
- Lower case letters used to denote pmf.
- The function:

$$p(x_i) = P(X = x_i) \quad \text{and} \quad \sum_{i=1}^{n} p(x_i) = 1$$

Cumulative Distribution Function (cdf)

- Sum successive probabilities.
- Gives the probability of X ≤ x_i
- Convenient, used in statistical table.
- Usually use upper case letters used to denote cdf.
- The function

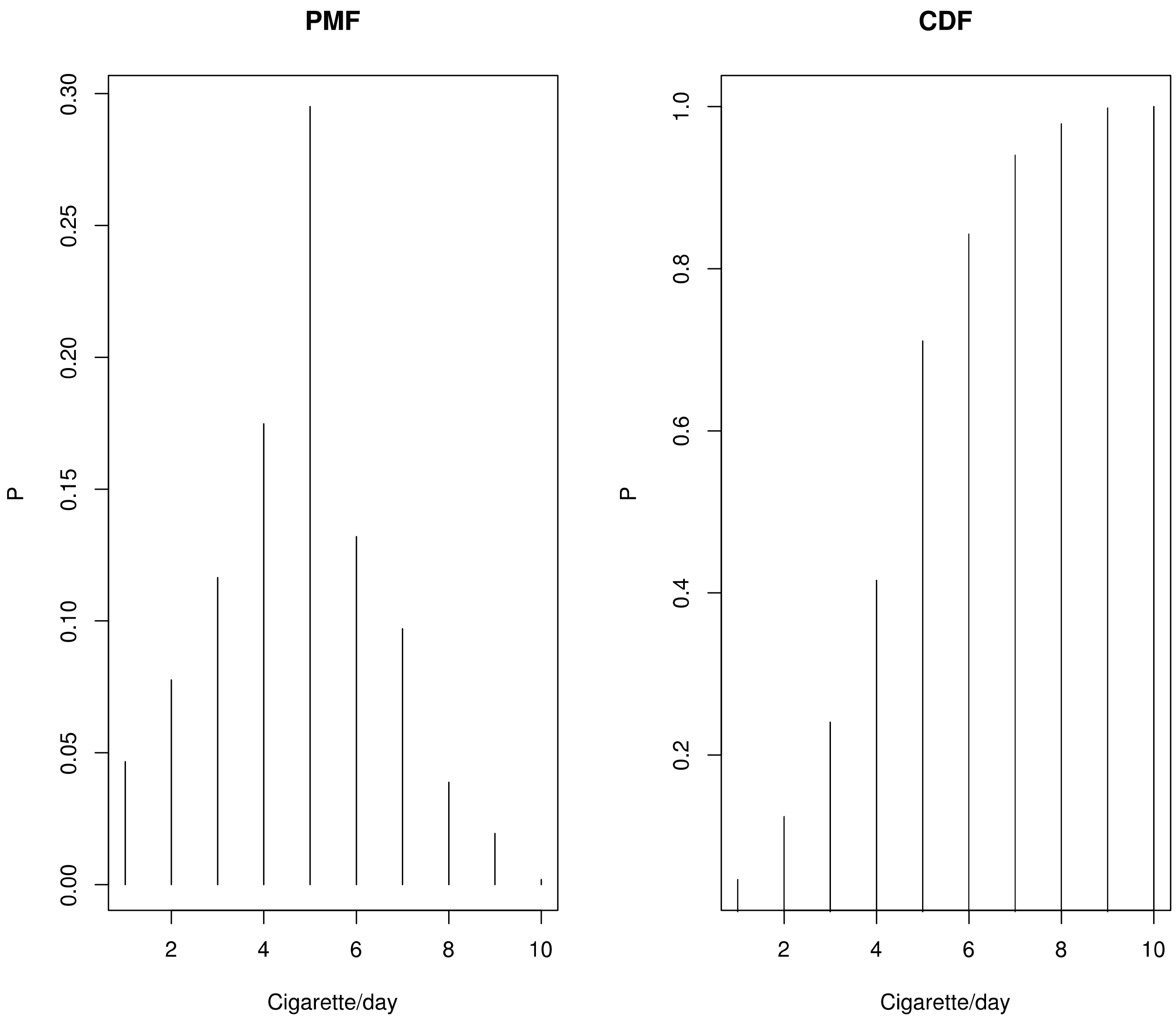
$$F(x) = P(X \leqslant x), \quad -\infty < x < \infty$$

- Non-decreasing and satisfies

$$\lim_{x \to -\infty} F(x) = 0 \quad \text{and} \quad \lim_{x \to \infty} F(x) = 1$$

Example 1:

Cigarette/day	Frequency	Probability	Cumulative Probability
1	120	0.0466	0.0466
2	200	0.0777	0.1243
3	300	0.1165	0.2408
4	450	0.1748	0.4155
5	760	0.2951	0.7107
6	340	0.1320	0.8427
7	250	0.0971	0.9398
8	100	0.0388	0.9786
9	50	0.0194	0.9981
10	5	0.0019	1.0000
Total	2575		



Bernoulli trial

- It is a trial that results in only one of two mutually exclusive outcomes, e.g. alive/dead, disease/no disease, HIV/no HIV, present/not present, success/failure etc.
- *Bernoulli process* is a sequence of Bernoulli trials with the following conditions:

  1. The outcome of each trial is one of two possible, mutually exclusive outcomes: success and failure.
  2. Probability of, let say, success p, remains constant from trial to trial. Probability of failure is given by 1 – p.
  3. The trials are independent: the outcome of any particular trial is unaffected by the knowledge of the outcome in another trial.

- Bernoulli random variable takes on only two values: 0 and 1, with probabilities 1 – p and p respectively. Thus the pmf:

$$
\begin{aligned}
&p(1)=p\\
&p(0)=1-p\\
&p(x)=0, \qquad \text{if } x\neq 0 \text{ and } x\neq 1
\end{aligned}
$$

alternatively

$$p(x)=\begin{cases} p^{x}(1-p)^{(1-x)}, & \text{if } x=0 \text{ or } x=1\\ 0, & \text{otherwise}\end{cases}$$

Binomial Distribution

- It is derived from Bernoulli trial.
- In n independent trials, n is fixed.
- Probability of success: p; probability of failure: 1 – p.
- Assumptions → *Bernoulli process*
- X is total number of success, with parameters n and p.
- Any particular sequence of *x* successes occurs with probability of $p^{x}(1-p)^{n-x}$ → multiplication principle.
- Total number of such sequences is $\binom{n}{x}$ ways to assign k successes to n trials.
- P(X=x) or p(x) is then the *number such sequences* times *probability of any particular sequence*.
- p(x) is given by

$$p(x)=\binom{n}{x}p^{x}(1-p)^{n-x}, \qquad x=0,1,2,\ldots,n$$

Properties

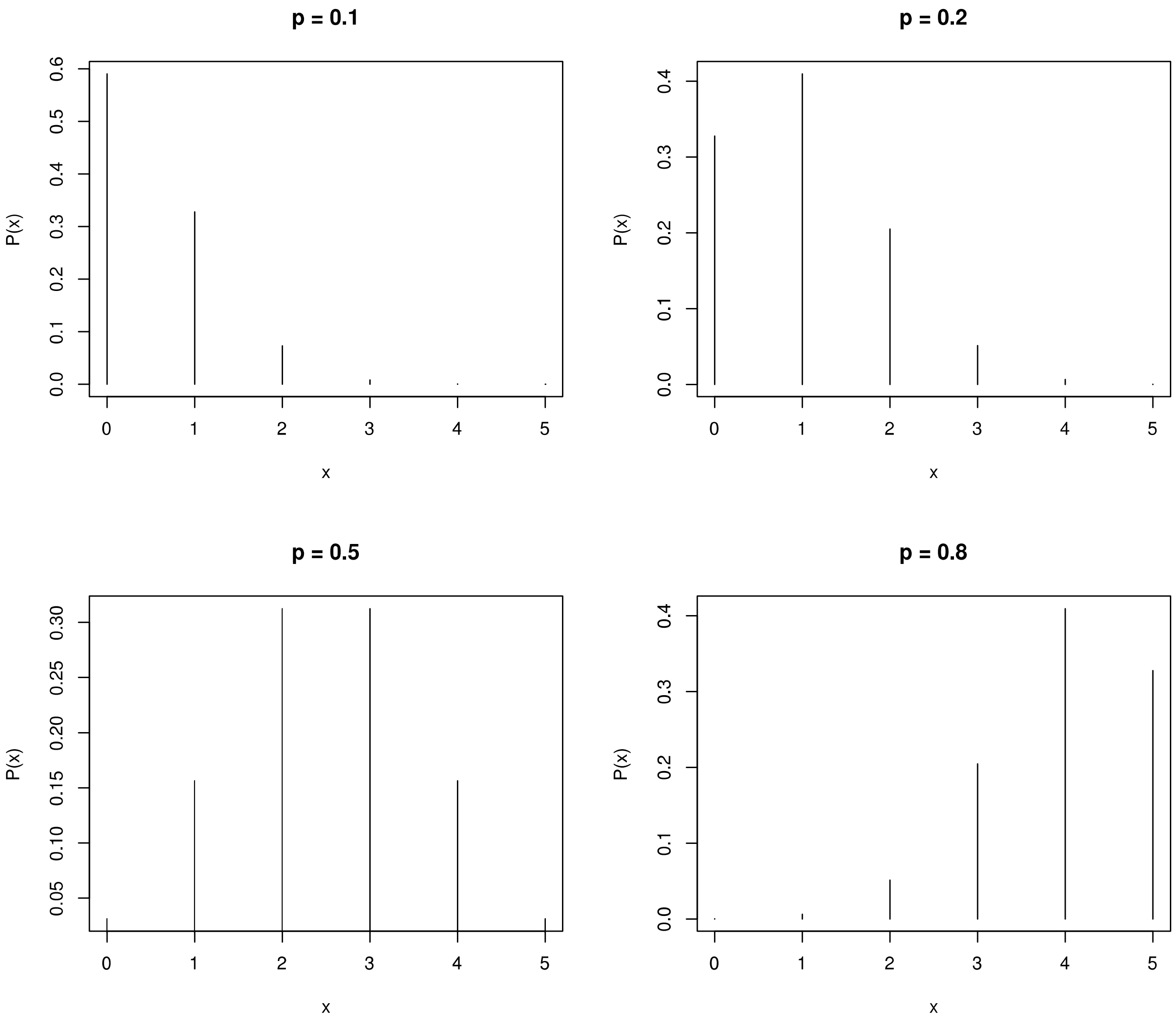
Mean = *np*

Variance = *np(1-p)*

Graphs of binomial probability mass function with different *p*

Using R

```
curve(dbinom(x, 5, 0.1), 0, 5, n = 6, type = "h")
curve(dbinom(x, 5, 0.2), 0, 5, n = 6, type = "h")
curve(dbinom(x, 5, 0.5), 0, 5, n = 6, type = "h")
curve(dbinom(x, 5, 0.8), 0, 5, n = 6, type = "h")
```



Example 2:

Probability of X = x

Suppose in Health Campus, USM it is known that 30% of the staffs are diabetic. If a random sample of 10 is selected among the staffs, what is the probability that exactly 4 staffs are diabetic?

$$P(X=4)=\binom{10}{4}0.3^{4}(0.7)^{6}=\frac{10!}{4!6!}(.0081)(.1176)=\frac{10.9.8.7}{4.3.2.1}(.0010)=.2100$$

Example 3:

Probability of X ≤ x

What is the probability that 2 or less staffs are diabetic?

$$\begin{aligned} P(X \leqslant 2) &= \sum_{x=0}^{2} \binom{10}{x} 0.3^{x} (0.7)^{10-x} \\ &= \binom{10}{0} 0.3^{0} (0.7)^{10} + \binom{10}{1} 0.3^{1} (0.7)^{9} + \binom{10}{2} 0.3^{2} (0.7)^{8} \\ &= .0282 + .1211 + .2355 \\ &= .3828 \end{aligned}$$

Example 4:

Probability of X > x

What is the probability that more than 2 staffs are diabetic?

$$P(X > 2) = 1 - P(X \leqslant 2) = 1 - .3828 = .6172$$

Using R

```
dbinom(x, n, p)
pbinom(q, n, p)
curve(dbinom(x, n, p), from, to, type = "h")

dbinom(4, 10, 0.3)  # Ex. 2
pbinom(2, 10, 0.3)  # Ex. 3
1 - pbinom(2, 10, 0.3)  # Ex. 4
curve(dbinom(x, 20, 0.3), 0, 20, n = 21, type = "h")
```

Using Statistical Table

Old style, please learn on your own.

Poisson Distribution

- It comes from binomial distribution i.e. limit of binomial distribution as the number of trials *n* approaches infinity and probability of success *p* approaches zero, so as *np* = λ.
- Assumptions → *Poisson process*:

1. The occurrences of the events are independent. The occurrence of an event in an interval of space or time does not affect the probability of second occurrence of the event in the same or different interval.
2. Infinite number of occurrences of the event is possible in the interval.
3. Probability of a single occurrence of the event in an interval is proportionate to interval length.
4. In a very small portion of the interval, probability of more than one occurrence of the event is negligible.

- p(x) is given by

$$p(x) = \frac{e^{-\lambda} \lambda^{x}}{x!}, \quad x = 0, 1, 2, \ldots; \quad \lambda > 0,$$

Properties

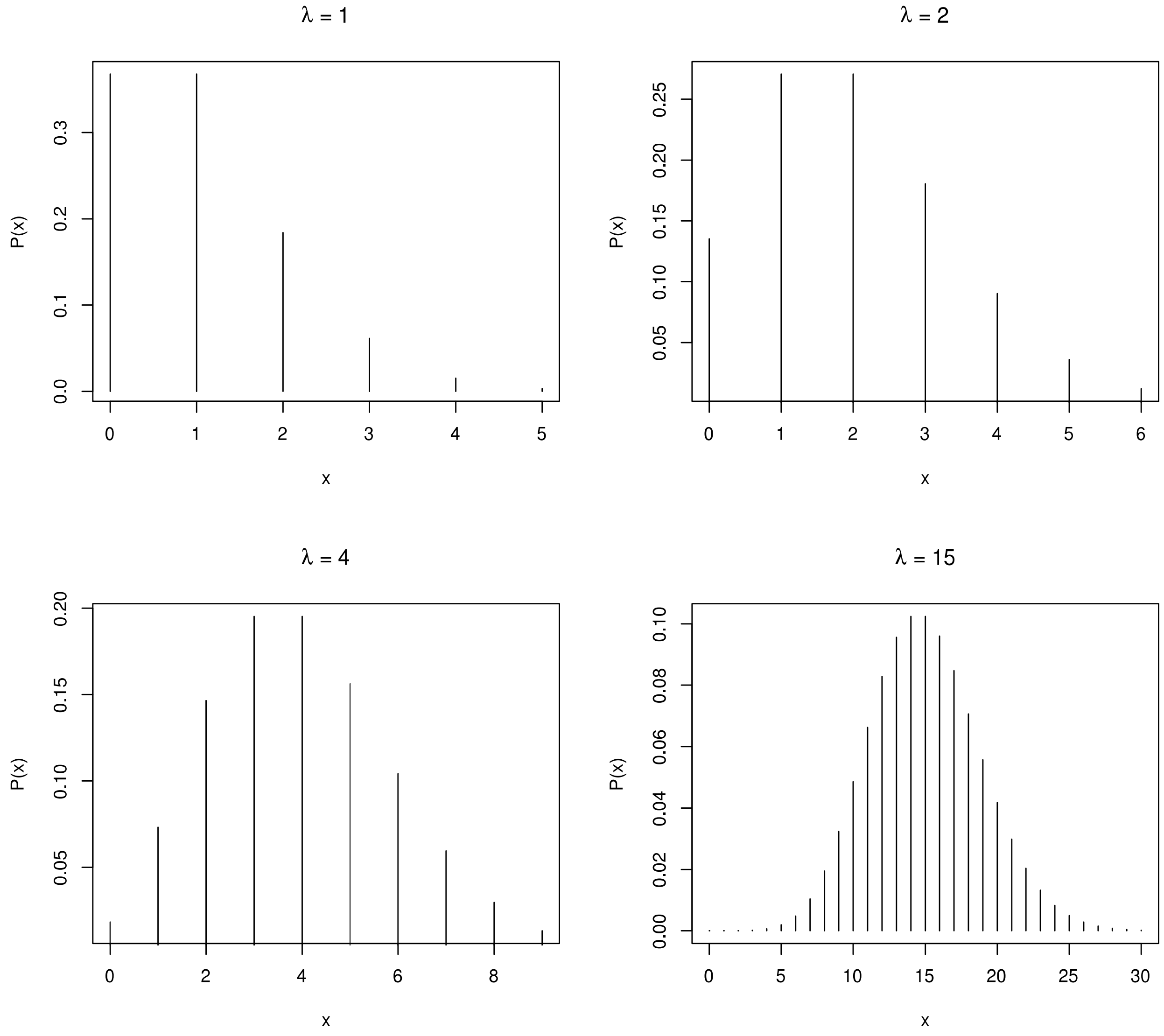
Mean = λ

Variance = λ

Graphs of poisson probability mass function with different *p*

<u>Using R</u>

```
curve(dpois(x, 1), 0, 5, n = 6, type = "h")
curve(dpois(x, 2), 0, 6, n = 7, type = "h")
curve(dpois(x, 4), 0, 9, n = 10, type = "h")
curve(dpois(x, 15), 0, 30, n = 31, type = "h")
```



λ = 1

λ = 2

λ = 4

λ = 15

Example 5:

Probability of X = x

Suppose the number of death due to motor vehicle accidents per day in Malaysia is on average 17.2 and it was found that the daily distribution follows Poisson distribution.

What is the probability that any randomly selected day will be the one with 10 death?

$$P(X=10)=\frac{e^{-17.2}17.2^{10}}{10!}=.0212$$

Example 6:

Probability of X ≤ x

What is the probability that any randomly selected day will be the one with less than 11 death?

$$P(X<11)=P(X\leqslant 10)=\sum_{x=0}^{10}\frac{e^{-17.2}17.2^{x}}{x!}=.0447$$

Example 7:

Probability of X >x

What is the probability that any randomly selected day will be more than 10 death?

$$P(X>10)=1-P(X\leqslant 10)=.9553$$

Using R

```
dpois(x, lambda)
ppois(q, lambda)
curve(dpois(x, lambda), from, to, type = "h")

dpois(10, 17.2)  # Ex. 6
ppois(10, 17.2)  # Ex. 7
1 - ppois(10, 17.2)  # Ex. 8
curve(dpois(x, 17.2), 0, 40, n = 41, type = "h")  # pmf
curve(ppois(x, 17.2), 0, 40, n = 41, type = "h")  # cdf
```

Using Statistical Table
Old style, please learn on your own.

Probability Distributions of Continuous Random Variables

Probability Distribution

- Known as **probability density function (pdf) –** for continuous random variable, replaces probability mass function for discrete random variable.
- If X is a continuous random variable with density function *f(x)*, with properties:

1. It is a nonnegative function,

$$f(x)\geqslant 0$$

2. Total area bounded by its curve and x-axis equal to 1,

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

3. For any two values *a* and *b*, in which *a* < *b*, the probability that *X* falls in the interval (*a*, *b*) is the area under the density function bounded by *a* and *b*,

$$P(a<X<b)=\int_a^b f(x)\,dx$$

4. Probability of any continuous random variable X taking on any particular value *a* is zero,

$$P(X=a)=\int_a^a f(x)\,dx=0$$

5. As such, if X is a continuous random variable, then,

$$P(a<X<b)=P(a\leqslant X<b)=P(a<X\leqslant b)=P(a\leqslant X\leqslant b)$$

Cumulative Distribution Function (cdf)

- For continuous random variable, the cdf,

$$F(x)=P(X\leqslant x)=\int_{-\infty}^x f(u)\,du,\quad -\infty<x<\infty$$

- Useful to know the probability that *X* falls in an interval,

$$P(a\leqslant X\leqslant b)=F(b)-F(a)$$

Normal Distribution

- Also known as Gaussian distribution.
- Probability density function, *f(x)* given by:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2},\quad -\infty<x<\infty$$

where $-\infty<\mu<\infty,\sigma>0,\ \mu-\text{mean},\sigma-\text{standard deviation}$

- Shorthand notation

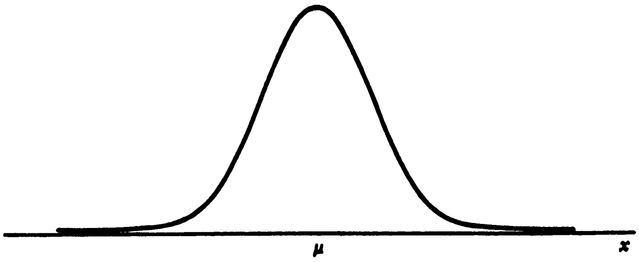
$$X\sim N(\mu,\sigma^2)$$

Properties

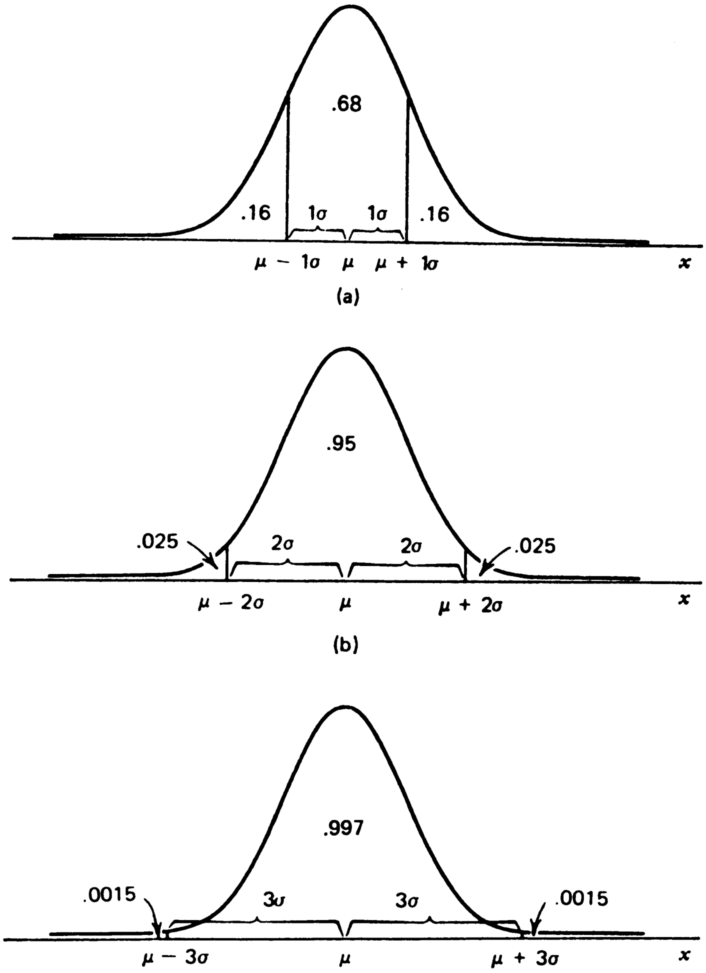
1. It is symmetrical about its mean, μ. The curve on either side of μ is a mirror image of the other side.
2. Its mean, median and mode are equal.
3. Total area bounded by its curve and x-axis equal to one square unit.
4. Areas bounded by 1σ, 2σ and 3σ from the mean in both direction are,

$$P(\mu-\sigma\leqslant X\leqslant\mu+\sigma)=0.6827$$
$$P(\mu-2\sigma\leqslant X\leqslant\mu+2\sigma)=0.9545$$
$$P(\mu-3\sigma\leqslant X\leqslant\mu+3\sigma)=0.9973$$

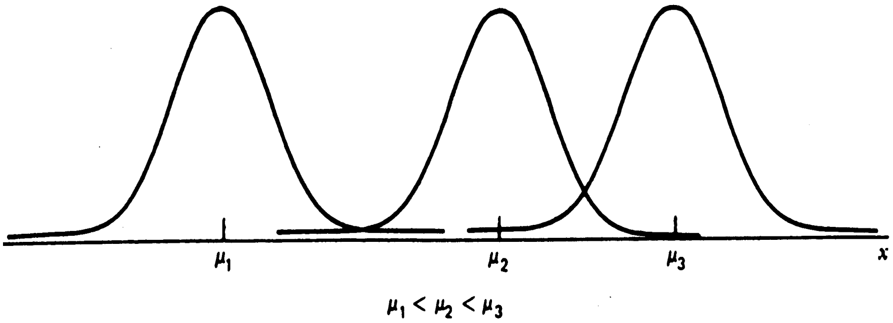
5. Different values of μ shift the graph along the x-axis.
6. Different values of σ affect the degree of peakness/flatness of the graph (kurtosis).



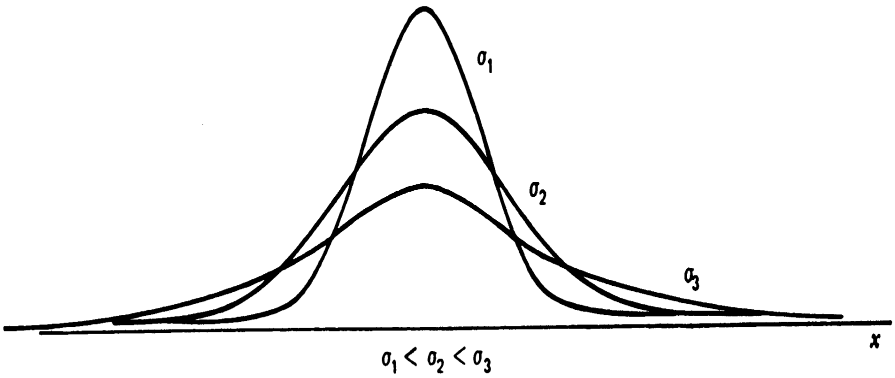
Graph of normal distribution.



Areas bounded by different σ.



Shift of graph with different μ.



Different degree of kurtosis by different σ.

*All these figures are from Daniel (1995)

Standard Normal Distribution

- Special form of normal distribution when $\mu=0, \sigma=1$ is called **standard normal distribution**.
- Standard normal pdf usually denoted by *f(z)* or small phi φ and its cdf by *F(z)* or large phi Φ (i.e if you read many different books, take note of the notation used).
- Standardize value of x by:

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

- Standard normal distribution is created by replacing *(x – μ)/σ* in function with *z,* thus the pdf *f(z),*

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

and its area,

$$P(a<Z<b)=\int_a^b \frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz$$

its cdf,

$$F(z)=P(Z\leqslant z)=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-w^2/2}dw$$

- Finding specific value/range from its probability → using inverse cdf denoted by $F^{-1}(p)$.

Properties
Mean = 0
Variance = 1

Example 8:

Given standard normal distribution, find the area between –1 < z < 1.

$$P(-1<Z<1)=P(Z<1)-P(Z<-1)=.8413-.1587=.6826$$

Example 9:

Given standard normal distribution, find the probability that a z value picked at random from population would be more than 1?

$$P(Z>1)=1-P(Z<1)=1-.8413=.1587$$

Example x:

Given standard normal distribution, find the probability that a z value picked at random from population would equal 1?

$$P(Z=1)=0\,!!!$$

Example 10:

Given following probabilities under standard normal distribution, find the value of a:

10(a) *P(Z < a)* = .8413

$$P(Z<a)=.8413$$
$$F^{-1}(P:Z<a)=F^{-1}(P=.8413)=a=.9998\approx 1$$

10(b) *P(Z>a)* = .0250

$$P(Z>a)=0.0250$$
$$1-P(Z<a)=0.0250$$
$$P(Z<a)=0.9750$$
$$F^{-1}[P(Z<a)]=F^{-1}(P=.9750)=a=1.9500$$

Example 11:

Suppose that the weights of individuals in a population is normally distributed with mean of 50kg and standard deviation of 10kg. What is the probability that the weight of a randomly picked person in the population falls between 40kg to 60kg?

$$P(40<X<60) \rightarrow \text{standardize to} \quad z=\frac{X-\mu}{\sigma}$$

$$P\left(\frac{40-50}{10}<Z<\frac{60-50}{10}\right)=P(-1<z<1)=.6826$$

Using R

```
pnorm(q, mean = 0, sd = 1)
qnorm(p, mean = 0, sd = 1)
curve(dnorm(x, mean, sd), from, to, type = “l”)

pnorm(1) - pnorm(-1)  # Ex. 8
1 - pnorm(1)  # Ex. 9
pnorm(1, lower.tail = FALSE)  # Ex. 9
qnorm(.8413)  # Ex. 10(a)
qnorm(0.975)  # Ex. 10(b)
pnorm(60, 50, 10) - pnorm(40, 50, 10)  # Ex. 11 = Ex. 8
curve(dnorm(x), -3, 3)  # standard normal
# the effect of different mean
curve(dnorm(x, 120, 15), 60, 180)
curve(dnorm(x, 140, 15), 60, 180, add = TRUE)
curve(dnorm(x, 100, 15), 60, 180, add = TRUE)
# the effect of different sigma
curve(dnorm(x, 120, 5), 60, 180)
curve(dnorm(x, 120, 10), 60, 180, add = TRUE)
curve(dnorm(x, 120, 15), 60, 180, add = TRUE)
```

Using Statistical Table
Old style, please learn on your own.

Topics for self-study

Mathematical expectation

- In relation to mean and variance.
- For discrete and continuous random variables.

Discrete distributions

- Geometric and negative binomial distributions.
- Hypergeometric distribution.

Continuous distributions

- Exponential distribution.
- Gamma distribution.
- Chi-square distribution.
- *t* distribution.
- *F* distribution.

References

Daniel, W. W. (1995). *Biostatistics: A foundation for analysis in the health sciences* (6th ed.). USA: John Wiley & Sons.

Rice, J. A. (1995). *Mathematical statistics and data analysis* (2nd ed.). USA: Duxbury Press.

Tijms, H. (2007). *Understanding probability: Chances rules in everyday life* (2nd ed.). New York, USA: Cambridge University Press.

Online resources

MIT’s Introduction to Probability and Statistics: https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/

PennState University’s Probability Theory and Mathematical Statistics: https://onlinecourses.science.psu.edu/stat414/

Wikipedia → Quite good for statistical distributions.