Chapter 4 Probability Distribution

Random Variables

- $0 \le P(X = x) \le 1$
- $\sum P(X = x) = 1$
- $E(X) \quad = \mu \quad = \sum x\, P(X = x)$
- $Var(X) = \sigma^2 \; = \sum (X - \mu)^2\, P(X = x)$
  $= E(X^2) - E(X)^2$

Question 1:
Given the following discrete probability distribution:

x	5	6	7	8
f(x)=P(X=x)	0.35	0.45	0.15	k

Find:

1. The value of k.

$$0.35 + 0.45 + 0.15 + k = 1 \implies \boxed{k = 0.05}$$

x	5	6	7	8
f(x)=P(X=x)	0.35	0.45	0.15	0.05

2. $P(X > 6) = 0.05 + 0.15 = 0.20$
3. $P(X \ge 6) = 0.45 + 0.15 + 0.05 = 0.65 \quad or \quad (1 - 0.35 = 0.65)$
4. $P(X < 4) = 0$
5. $P(X > 3) = 1$

Question 2:
Which of the following functions can be a probability distribution of a discrete random variable?

(a)

x	g(x)
0	0.6
1	-0.2
2	0.5
3	0.1
✗	

(b)

x	g(x)
0	0.4
1	0.1
2	0.5
3	0.2
✗	

(c)

x	g(x)
0	0.1
1	1.2
2	-0.6
3	0.3
✗	

(d)

x	g(x)
0	0.3
1	0.1
2	0.5
3	0.1
✓	

(e)

x	g(x)
0	0.2
1	0.4
2	0.3
3	0.4
✗	

(f)

x	g(x)
0	0.1
1	0.2
2	0.3
3	0.1
✗	

Question 3:
Which of the following is a probability distribution function:

$a.\ f(x) = \frac{x+1}{10}\ ;\ x = 0,1,2,3,4$

$b.\ f(x) = \frac{x-1}{5}\ ;\ x = 0,1,2,3,4$

$c.\ f(x) = \frac{1}{5}\ ;\ x = 0,1,2,3,4$

$d.\ f(x) = \frac{5-x^2}{6}\ ;\ x = 0,1,2,3$

a.

$$f(x) = \frac{x+1}{10};\ \ x = 0, 1, 2, 3, 4$$

x	0	1	2	3	4
$f(x)$	$^1/_{10}$	$^2/_{10}$	$^3/_{10}$	$^4/_{10}$	$^5/_{10}$

$f(x)$ *is not a P.D.F because* $\sum f(x) \neq 1$

b.

$$f(x) = \frac{x-1}{5};\ \ x = 0, 1, 2, 3, 4$$

x	0	1	2	3	4
$f(x)$	$-^1/_5$				

$f(x)$ *is not a P.D.F because every* $f(x)$ *shoud be* $0 \leq f(x) \leq 1$

c.

$$f(x) = \frac{1}{5};\ \ x = 0, 1, 2, 3, 4$$

x	0	1	2	3	4
$f(x)$	$^1/_5$	$^1/_5$	$^1/_5$	$^1/_5$	$^1/_5$

$f(x)$ *is a P.D.F*

d.

$$f(x) = \frac{5-x^2}{6};\ \ x = 0, 1, 2, 3$$

x	0	1	2	3
$f(x)$				$-^4/_6$

$f(x)$ *is not a P.D.F because every* $f(x)$ *shoud be* $0 \leq f(x) \leq 1$

Question 4:

Given the following discrete probability distribution:

x	5	6	7	8
f(x)=P(X=x)	2k	3k	4k	k

Find the value of k.

$$2k + 3k + 4k + k = 1$$

$$10k = 1 \implies \boxed{k = 0.1}$$

x	5	6	7	8
f(x)=P(X=x)	0.2	0.3	0.4	0.1

Question 5:

Let X be a discrete random variable with probability mass function:
$f(x) = cx$; $x = 1,2,3,4$ What is the value of c?

x	1	2	3	4
$P(X = x)$	c	2c	3c	4c

$$c + 2c + 3c + 4c = 1 \quad \implies c = \frac{1}{10}$$

Then probability mass function:

x	1	2	3	4
$P(X = x)$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{4}{10}$

Question 6:

Let X be a discrete random variable with probability function given by:

$$f(x) = c(x^2 + 2) \quad ; \quad x = 0,1,2,3$$

$f(0) = c(0^2 + 2) = 2c$

$f(1) = c(1^2 + 2) = 3c$

$f(2) = c(2^2 + 2) = 6c$

$f(3) = c(3^2 + 2) = 11c$

x	0	1	2	3
f(x)	2c	3c	6c	11c

$$2c + 3c + 6c + 11c = 1 \qquad c = \frac{1}{22} = 0.04545$$

x	0	1	2	3
f(x)	$\frac{2}{22}$	$\frac{3}{22}$	$\frac{6}{22}$	$\frac{11}{22}$

Question 7:

Given the following discrete probability distribution:

x	5	6	7	8
f(x)=P(X=x)	0.2	0.4	0.3	0.1

Find:

1. Find the mean of the distribution $\mu = \mu_X = E(X)$.

$$E(X) = \mu = \mu_X = \sum_{x=5}^{8} x\,P(X = x)$$
$$= (5)(0.2) + (6)(0.4) + (7)(0.3) + (8)(0.1) = 6.3$$

2. Find the variance of the distribution $\sigma^2 = \sigma_X^2 = \mathrm{Var}(X)$.

$$E(X^2) = (5^2 \times 0.2) + (6^2 \times 0.4) + (7^2 \times 0.3) + (8^2 \times 0.1) = 40.5$$

$$\mathrm{Var}(X) = E(X^2) - E(X)^2$$
$$= 40.5 - 6.3^2 = 0.81$$

Or:

$$\mathrm{Var}(X) = \sigma^2 = \sigma_X^2 = \sum (x - \mu)^2\,P(X = x)$$
$$= \sum_{x=5}^{8} (x - 6.3)^2\,P(X = x)$$

$$= (5 - 6.3)^2(0.2) + (6 - 6.3)^2(0.4) + (7 - 6.3)^2(0.3) + (8 - 6.3)^2(0.1) = 0.81$$

Question 8:

Given the following discrete distribution:

x	-1	0	1	2	3	4
P(X=x)	0.15	0.30	M	0.15	0.10	0.10

1. The value of M is equal to

$$\boxed{M = 1 - (0.15 + 0.30 + 0.15 + 0.10 + 0.10) = 1 - 0.80 = 0.20}$$

2. P (X $\le$0.5) = $\boxed{0.15 + 0.30 = 0.45}$

3. P(X=0) = $\boxed{0.30}$

4. The expected (mean) value E[X] is equal to

$$\boxed{E(X) = (-1 \times 0.15) + (0 \times 0.30) + (1 \times 0.20) + (2 \times 0.15) + (3 \times 0.10) + (4 \times 0.10) = 1.05}$$

<u>Question 9:</u>

The average length of stay in a hospital is useful for planning purposes. Suppose that the following is the probability distribution of the length of stay (X) in a hospital after a minor operation:

Length of stay (days)	3	4	5	6
Probability	0.4	0.2	0.1	k

(1) The value of k is

$$\boxed{k = 1 - (0.4 + 0.2 + 0.1) = 1 - 0.7 = 0.3}$$

(2) $P(X \le 0) =$

$$\boxed{0}$$

(3) $P(0 < X \le 5) =$

$$\boxed{0.4 + 0.2 + 0.1 = 0.7}$$

(4) $P(X \le 5.5) =$

$$\boxed{0.4 + 0.2 + 0.1 = 0.7}$$

(5) The probability that the patient will stay at most 4 days in a hospital after a minor operation is equal to

$$\boxed{0.4 + 0.2 = 0.6}$$

(6) The average length of stay in a hospital is

$$\boxed{E(X) = (3 \times 0.4) + (4 \times 0.2) + (5 \times 0.1) + (6 \times 0.3) = 4.3}$$

<u>Question 10</u>:

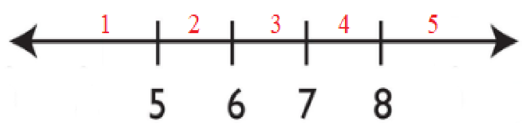
Given the following discrete probability distribution:

x	5	6	7	8
f(x)=P(X=x)	0.2	0.4	0.3	0.1

1. Find the cumulative distribution of X.

x	5	6	7	8
$F(x) = P(X \leq x)$	0.2	0.6	0.9	1

$$F(x) = \begin{cases} 0 & X < 5 \\ 0.2 & 5 \leq X < 6 \\ 0.6 & 6 \leq X < 7 \\ 0.9 & 7 \leq X < 8 \\ 1 & X \geq 8 \end{cases}$$



2. From the cumulative distribution of X, find:

a) $P(X \leq 7) \quad = 0.9$

b) $P(X \leq 6.5) = P(X \leq 6)$ =0.6

c) $P(X > 6) \quad = 1 - P(X \leq 6) = 1 - 0.6 = 0.4$

d) $P(X > 7) \quad = 1 - P(X \leq 7) = 1 - 0.9 = 0.1$

<u>Question 11</u>:

Given that the cumulative distribution of random variable T, is:

$$F(t) = P(T \leq t) = \begin{cases} 0 & t < 1 \\ 1/2 & 1 \leq t < 3 \\ 8/12 & 3 \leq t < 5 \\ 3/4 & 5 \leq t < 7 \\ 1 & t \geq 7 \end{cases}$$

1. Find $P(T = 5)$

T	1	3	5	7
f(t)	$\frac{1}{2} - 0 = \mathbf{0.5}$	$\frac{8}{12} - \frac{1}{2} = \mathbf{0.167}$	$\frac{3}{4} - \frac{8}{12} = \mathbf{0.083}$	$1 - \frac{3}{4} = \mathbf{0.25}$

$$P(T = 5) = 0.083$$

2. Find $P(1.4 < T < 6) = 0.167 + 0.083 = 0.25$

Binomial Distribution:

$$P(X=x)=\binom{n}{x}p^x q^{n-x} \;\; ; \quad x=0,1\ldots.,n$$

$$*\,E(X)=np \qquad *\,Var(X)=npq$$

$$q=1-p$$

Question 1:

Suppose that 25% of the people in a certain large population have high blood pressure. A Sample of 7 people is selected at random from this population. Let X be the number of people in the sample who have high blood pressure, follows a binomial distribution then

1) The values of the parameters of the distribution are*:*

$$\boxed{p=0.25 \quad , \quad n=7}$$

A	7 , 0.75	B	7 , 0.25	C	0.25 , 0.75	D	25 , 7

2) The probability that we find exactly one person with high blood pressure, is:

X	0	1	2	3	4	5	6	7
$P(X=x)$		*						

$$\boxed{P(X=1)=\binom{7}{1}(0.25)^1(0.75)^6=0.31146}$$

3) The probability that there will be at most one person with high blood pressure, is:

X	0	1	2	3	4	5	6	7
$P(X=x)$	*	*						

$$\boxed{P(X\le 1)=\binom{7}{0}(0.25)^0(0.75)^7+\binom{7}{1}(0.25)^1(0.75)^6=0.4449}$$

4) The probability that we find more than one person with high blood pressure, is:

X	0	1	2	3	4	5	6	7
$P(X=x)$			*	*	*	*	*	*

$$\boxed{P(X>1)=1-P(X\le 1)=1-0.4449=0.5551}$$

Question 2:

In some population it was found that the percentage of adults who have hypertension is 24 percent. Suppose we select a simple random sample of five adults from this population. Then the probability that the number of people who have hypertension in this sample, will be:

$$\boxed{p = 0.24 \quad , \quad n = 5}$$

1. Zero*:*

$$\boxed{P(X = 0) = \binom{5}{0}(0.24)^0(0.76)^5 = 0.2536}$$

2. Exactly one

$$\boxed{P(X = 1) = \binom{5}{1}(0.24)^1(0.76)^4 = 0.4003}$$

3. Between one and three, inclusive

$$\boxed{P(1 \le X \le 3) = \binom{5}{1}(0.24)^1(0.76)^4 + \binom{5}{2}(0.24)^2(0.76)^3 + \binom{5}{3}(0.24)^3(0.76)^2 = 0.7330}$$

4. Two or fewer (at most two):

$$\boxed{P(X \le 2) = \binom{5}{0}(0.24)^0(0.76)^5 + \binom{5}{1}(0.24)^1(0.76)^4 + \binom{5}{2}(0.24)^2(0.76)^3 = 0.9067}$$

5. Five:

$$\boxed{P(X = 5) = \binom{5}{5}(0.24)^5(0.76)^0 = 0.0008}$$

6. The mean of the number of people who have hypertension is equal to:

$$\boxed{E(X) = np = 5 \times 0.24 = 1.2}$$

7. The variance of the number of people who have hypertension is:

$$\boxed{Var(X) = npq = 5 \times 0.24 \times 0.76 = 0.912}$$

More Exercises

Exercise 1:

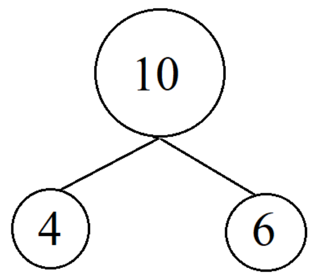
Find:

1. $6! \quad = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$
2. ${}_8C_3 \quad = \frac{8!}{3!\,(8-3)!} = \frac{8!}{3! \times 5!} = \quad 56$
3. ${}_8C_{10} \quad = 0$
4. ${}_8C_{-5} \quad = 0$

Exercise 2:

A box contains 10 cards numbered from 1 to 10. In how many ways can we select 4 cards out of this box?

$$\text{Answer } = {}_{10}C_4 \; = \frac{10!}{4!\,(10-4)!} = \frac{10!}{4! \times 6!}$$
$$= \frac{10\times9\times8\times7\times6!}{(4\times3\times2\times1)\;6!} = \frac{10\times9\times8\times7}{4\times3\times2\times1}$$
$$= \; 210$$



Exercise 3:

The manager of a certain bank has recently examined the credit card account balances for the customers of his bank and found that 20% of the customers have excellent records. Suppose that the manager randomly selects a sample of 4 customers.

(A) Define the random variable X as:

X = The number of customers in the sample having excellent records.

Find the probability distribution of X.

$X \sim Binomial\;(n\;,p\;)$

$n = 4 \quad (Number\ of\ trials)$

$p = \frac{20}{100} = 0.2 \quad (Probability\ of\ success)$

$q = 1 - p = 1 - 0.2 = 0.8 \quad (Probability\ of\ failure)$

$x = 0, 1, 2, 3, 4 \quad (Possible\ values\ of\ X)$

(a) The probability function in a mathematical formula:

$$P(X=x)=\begin{cases}\dfrac{n!}{x!\,(n-x)!}\;p^{x}\,q^{n-x} & ;\quad x=0,1,2,\dots,n\\ 0 & ;Otherwise\end{cases}$$

$$P(X=x)=\begin{cases}\dfrac{4!}{x!\,(4-x)!}\;(0.2)^{x}\,(0.8)^{4-x} & ;\quad x=0,1,2,3,4\\ 0 & ;Otherwise\end{cases}$$

(b) The probability function in a table:

x	$P(X=x)$
0	$\dfrac{4!}{0!\,(4-0)!}\;(0.2)^{0}\,(0.8)^{4-0}=(1)(0.2)^{0}(0.8)^{4}=0.4096$
1	$\dfrac{4!}{1!\,(4-1)!}\;(0.2)^{1}\,(0.8)^{4-1}=(4)(0.2)^{1}(0.8)^{3}=0.4096$
2	$\dfrac{4!}{2!\,(4-2)!}\;(0.2)^{2}\,(0.8)^{4-2}=(6)(0.2)^{2}\,(0.8)^{2}=0.1536$
3	$\dfrac{4!}{3!\,(4-3)!}\;(0.2)^{3}\,(0.8)^{4-3}=(4)(0.2)^{3}\,(0.8)^{1}=0.0256$
4	$\dfrac{4!}{4!\,(4-4)!}\;(0.2)^{4}\,(0.8)^{4-4}=(1)(0.2)^{4}\,(0.8)^{0}=0.0016$
	Total = 1

x	$P(X=x)$
0	0.4096
1	0.4096
2	0.1536
3	0.0256
4	0.0016

(B) Find:

1. The probability that there will be 3 customers in the sample having excellent records.
$$P(X = 3) = 0.0256$$
2. The probability that there will be no customers in the sample having excellent records.
$$P(X = 0) = 0.4096$$
3. The probability that there will be at least 3 customers in the sample having excellent records.
$$P(X \geq 3) = P(x = 3) + P(X = 4) = 0.0256 + 0.0016$$
$$= 0.0272$$
4. The probability that there will be at most 2 customers in the sample having excellent records.
$$P(X \leq 2) = P(x = 0) + P(X = 1) + P(X = 2)$$
$$= 0.4096 + 0.4096 + 0.1536$$
$$= 0.9728$$
5. The expected number of customers having excellent records in the sample.
$$E(X) = \mu = \mu_X = np = 4 \times 0.2 = 0.8$$
6. The variance of the number of customers having excellent records in the sample.
$$Var(X) = \sigma^2 = \sigma_X^2 = npq = 4 \times 0.2 \times 0.8 = 0.64$$

<u>Exercise 4</u>: (Do it at home for yourself)
In a certain hospital, the medical records show that the percentage of lung cancer patients who smoke is 75%. Suppose that a doctor randomly selects a sample of 5 records of lung cancer patients from this hospital.
(A) Define the random variable X as:
X = The number of smokers in the sample.
Find the probability distribution of X.

(B) Find:

1. The probability that there will be 4 smokers in the sample.
2. The probability that there will be no smoker in the sample.
3. The probability that there will be at least 2 smokers in the sample.
4. The probability that there will be at most 3 smokers in the sample.
5. The expected number of smokers in the sample.
6. The variance of the number of smokers in the sample.

Poisson distribution:

$$P(X=x)=\frac{e^{-\lambda}\ \lambda^{x}}{x!}\quad ;\quad x=0,1,2,\dots$$
$$E(X)=Var(X)=\lambda$$

Question 1:

The number of serious cases coming to a hospital during a night follows a Poisson distribution with an average of 10 persons per night, then:

1) The probability that 12 serious cases coming in the next night, is:

$$\boxed{\begin{array}{c}\lambda_{one\ night}=10\\ \\ P(X=12)=\dfrac{e^{-10}\ 10^{12}}{12!}=0.09478\end{array}}$$

2) The average number of serious cases in a two nights' period is:

$$\boxed{\lambda_{two\ nights}=20}$$

3) The probability that 20 serious cases coming in next two nights is:

$$\boxed{\begin{array}{c}\lambda_{two\ nights}=20\\ \\ P(X=20)=\dfrac{e^{-20}\ 20^{20}}{20!}=0.0888\end{array}}$$

<u>Question 2:</u>

Given the mean number of serious accidents per year in a large factory is five. If the number of accidents follows a Poisson distribution, then the probability that in the next year there will be:

1. Exactly seven accidents:

$$\lambda_{one\ year} = 5$$

$$P(X = 7) = \frac{e^{-5}\ 5^7}{7!} = 0.1044$$

2. No accidents

$$P(X = 0) = \frac{e^{-5}\ 5^0}{0!} = 0.0067$$

3. one or more accidents

X	0	1	2	3	4	5	6	...
$P(X = x)$		*	*	*	*	*	*	*

$$P(X \geq 1) = 1 - P(X < 1)$$
$$= 1 - P(X = 0)$$
$$= 1 - 0.0067 = 0.9933$$

4. The expected number (mean) of serious accidents in the next two years is equal to

$$\lambda_{two\ years} = 10$$

5. The probability that in the next two years there will be three accidents

$$\lambda_{two\ years} = 10$$

$$P(X = 3) = \frac{e^{-10}\ 10^3}{3!} = 0.0076$$

More Exercise

Exercise 1:

Suppose that in a certain city, the weekly number of infected cases with Corona virus (COVID-19) has a Poisson distribution with an average (mean) of 5 cases per week.

(A) Find:

1. The probability distribution of the weekly number of infected cases (X).

$$P(X=x)=\begin{cases}\dfrac{e^{-\lambda}\,\lambda^{x}}{x!} & ;\quad x=0,1,2,3,\dots\\ 0 & ;Otherwise\end{cases}$$

$$\lambda=5$$

$$P(X=x)=\begin{cases}\dfrac{e^{-5}\,5^{x}}{x!} & ;\quad x=0,1,2,3,\dots\\ 0 & ;Otherwise\end{cases}$$

2. The probability that there will be 2 infected cases this week.

$$P(X=2)=\frac{e^{-5}\,5^{2}}{2!}=0.0842$$

3. The probability that there will be 1 infected case this week.

$$P(X=1)=\frac{e^{-5}\,5^{1}}{1!}=0.0337$$

4. The probability that there will be no infected cases this week.

$$P(X=0)=\frac{e^{-5}\,5^{0}}{0!}=0.0067$$

5. The probability that there will be at least 3 infected cases this week.

$$\begin{aligned}P(X\geq 3)&=1-P(X<3)=1-P(X\leq 2)\\&=1-[P(X=0)+P(X=1)+P(X=2)]\\&=1-[\,0.0067+0.0337+0.0842\,]\\&=1-0.1246=0.8754\end{aligned}$$

6. The probability that there will be at most 2 infected cases this week.

$$\begin{aligned}P(X\leq 2)&=P(X=0)+P(X=1)+P(X=2)\\&=0.0067+0.0337+0.0842\\&=0.1246\end{aligned}$$

7. The expected number (mean/average) of infected cases this week.

$$E(X) = \mu = \mu_X = \lambda = 5$$

8. The variance of the number of infected cases this week.

$$Var(X) = \sigma^2 = \sigma_X^2 = \lambda = 5$$

(B): Find:

1. The average (mean) of the number infected cases in a day.

$$\lambda = \frac{5}{7} = 0.7143$$

2. The probability distribution of the daily number of infected cases (X).

$$P(X = x) = \begin{cases} \dfrac{e^{-\lambda}\ \lambda^x}{x!} & ; \quad x = 0, 1, 2, 3, \ldots \\ 0 & ; Otherwise \end{cases}$$

$$\lambda = \frac{5}{7}$$

$$P(X = x) = \begin{cases} \dfrac{e^{-\frac{5}{7}}\ \left(\frac{5}{7}\right)^x}{x!} & ; \quad x = 0, 1, 2, 3, \ldots \\ 0 & ; Otherwise \end{cases}$$

3. The probability that there will be 2 infected cases tomorrow.

$$P(X = 2) = \frac{e^{-\frac{5}{7}}\ \left(\frac{5}{7}\right)^2}{2!} = 0.1249$$

4. The probability that there will be 1 infected case tomorrow.

$$P(X = 1) = \frac{e^{-\frac{5}{7}}\ \left(\frac{5}{7}\right)^1}{1!} = 0.3497$$

5. The probability that there will be no infected cases tomorrow.

$$P(X = 0) = \frac{e^{-\frac{5}{7}}\ \left(\frac{5}{7}\right)^0}{0!} = 0.4895$$

6. The probability that there will be at most 2 infected cases tomorrow.

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$
$$= 0.4895 + 0.3497 + 0.1249$$
$$= 0.9641$$

7. The probability that there will be at least 2 infected cases tomorrow.

$$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X \leq 1)$$
$$= 1 - [P(X = 0) + P(X = 1)]$$
$$= 1 - [0.4895 + 0.3497]$$
$$= 1 - 0.8392 = 0.1608$$

8. The expected number (mean/average) of infected cases tomorrow.

$$E(X) = \mu = \mu_X = \lambda = \frac{5}{7} = 0.7143$$

9. The variance of the number of infected cases tomorrow.

$$Var(X) = \sigma^2 = \sigma_X^2 = \lambda = \frac{5}{7} = 0.7143$$

(C): Assuming that 4 weeks are in a month, find:

1. The average (mean) of the number infected cases per month.

$$E(X) = \mu = \mu_X = \lambda = 5 \times 4 = 20$$

2. The variance of the number of infected cases per month.

$$Var(X) = \sigma^2 = \sigma_X^2 = \lambda = 5 \times 4 = 20$$

The Normal Distribution:

μ

Mean = Median = Mode

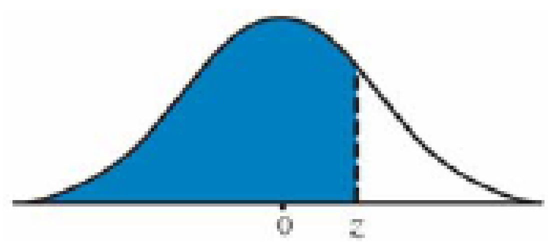
Normal distribution $X \sim N(\mu\,, \sigma^2)$
Standard normal $Z \sim N(0\;, 1)$

Question 1:

Given the standard normal distribution, Z~N (0, 1), find:

1. $P(Z < 1.43) = 0.92364$

z	0.00	0.01	0.02	0.03	0.04
0.00	0.50000	0.50399	0.50798	0.5[illegible]197	0.51595
0.10	0.53983	0.54380	0.54776	0.5[illegible]172	0.55567
0.20	0.57926	0.58317	0.58706	0.5[illegible]095	0.59483
0.30	0.61791	0.62172	0.62552	0.6[illegible]930	0.63307
0.40	0.65542	0.65910	0.66276	0.6[illegible]640	0.67003
0.50	0.69146	0.69497	0.69847	0.7[illegible]194	0.70540
0.60	0.72575	0.72907	0.73237	0.7[illegible]565	0.73891
0.70	0.75804	0.76115	0.76424	0.7[illegible]730	0.77035
0.80	0.78814	0.79103	0.79389	0.7[illegible]673	0.79955
0.90	0.81594	0.81859	0.82121	0.8[illegible]381	0.82639
1.00	0.84134	0.84375	0.84614	0.8[illegible]849	0.85083
1.10	0.86433	0.86650	0.86864	0.8[illegible]076	0.87286
1.20	0.88493	0.88686	0.88877	0.8[illegible]065	0.89251
1.30	0.90320	0.90490	0.90658	0.9[illegible]824	0.90988
1.40	[illegible]	[illegible]	[illegible]	0.92364	0.92507
1.50	0.93319	0.93448	0.93574	0.93699	0.93822
1.60	0.94520	0.94630	0.94738	0.94845	0.94950
1.70	0.95543	0.95637	0.95728	0.95818	0.95907
1.80	0.96407	0.96485	0.96562	0.96638	0.96712



$P\left(Z < \boxed{\text{أطراف الجدول}}\right) = \boxed{\text{داخل الجدول}}$

2. $P(Z > 1.67) = 1 - P(Z < 1.67) = 1 - 0.95254 = 0.04746$

3. $P(-2.16 < Z < -0.65)$

$= P(Z < -0.65) - P(Z < -2.16)$
$= \quad 0.25785 \quad - \quad 0.01539 \quad = 0.24246$

Question 2:

Given the standard normal distribution, Z~N(0,1), find:

1. $P(Z > 2.71) = 1 - P(Z < 2.71) = 1 - 0.99664 = 0.00336$

2. $P(-1.96 < Z < 1.96)$

$$= P(Z < 1.96) - P(Z < -1.96)$$
$$= \quad 0.9750 \quad - \quad 0.0250 \quad = 0.9500$$

3. $P(Z = 1.33) = 0$
4. $P(Z = 0.67) = 0$

5. If $P(Z < a) = 0.99290$, then the value of $a = 2.45$
6. If $P(Z < a) = 0.62930$, then the value of $a = 0.33$

7. If $P(Z > a) = 0.63307 \Longrightarrow P(Z < a) = 1 - 0.63307$

$$\Longrightarrow P(Z < a) = 0.36693 \Longrightarrow a = -0.34$$

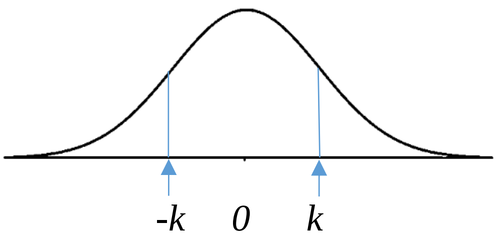
8. If $P(Z > a) = 0.02500 \Longrightarrow P(Z < a) = 1 - 0.02500$

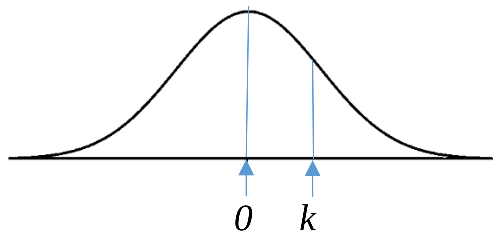
$$\Longrightarrow P(Z < a) = 0.97500 \Longrightarrow a = 1.96$$

9. $Z_{0.9750} = 1.96$
10. $Z_{0.0392} = -1.76$
11. $Z_{0.01130} = -2.28$
12. $Z_{0.99940} = 3.24$
13. If $Z_{0.08} = -1.40$ then the value of $Z_{0.92}$ equals to:

A	-1.954	B	1	C	1.40	D	-1.40

14.If $P(-k < Z < k) = 0.8132$, then the value of k =





$$\Longrightarrow 2 \times P(0 < Z < k) = 0.8132$$

$$\Longrightarrow \quad P(0 < Z < k) = 0.4066$$

$$\Longrightarrow P(Z < k) - P(Z < 0) = 0.4066$$

$$\Longrightarrow P(Z < k) - \quad 0.5 \quad = 0.4066$$

$$\Longrightarrow P(Z < k) = 0.9066$$

$$\Longrightarrow k = 1.32$$

<u>Question 3:</u>

Given the standard normal distribution, then:

1) $P(-1.1 < Z < 1.1)$ *is:*

$$\Longrightarrow P(Z < 1.1) - P(Z < -1.1)$$

$$0.86433 - 0.13567 = 0.72866$$

2) $P(Z > -0.15)$ *is:*

$$= 1 - P(Z < -0.15)$$

$$= 1 - 0.44038 = 0.55962$$

3) The k value that has an area of 0.883 to its right, is:

Left	*Right*
<	>

$$P(Z > k) = 0.883$$
$$P(Z < k) = 1 - 0.883$$
$$P(Z < k) = 0.117$$
$$k = -1.19$$

Question 4:
The finished inside diameter of a piston ring is normally distributed with a mean 12 cm and standard deviation of 0.03 cm. Then,

1. The proportion of rings that will have inside dimeter less than 12.05.

$$X \sim N(\mu \ ,\sigma^2)$$
$$X \sim N(12 \ ,0.03^2)$$

$$P(X < 12.05) = P\left(Z < \frac{12.05 - \mu}{\sigma}\right)$$

$$= P\left(Z < \frac{12.05 - 12}{0.03}\right) = P(Z < 1.67) = 0.9525$$

2. The proportion of rings that will have inside dimeter exceeding 11.97.

$$P(X > 11.97) = P\left(Z > \frac{11.97 - \mu}{\sigma}\right)$$

$$= P\left(Z > \frac{11.97 - 12}{0.03}\right) = P(Z > -1)$$
$$= 1 - P(Z < -1)$$
$$= 1 - 0.1587 = 0.8413$$

3. The proportion of rings that will have inside dimeter between 11.95 and 12.05.

$$P(11.95 < X < 12.05) = P\left(\frac{11.95 - 12}{0.03} < Z < \frac{12.05 - 12}{0.03}\right)$$

$$= P(\ -1.67 \quad < Z < \quad 1.67)$$

$$= P(Z < 1.67) - P(Z < -1.67)$$

$$= \quad 0.9525 \quad - \quad 0.0475 = 0.905$$

<u>Question 5:</u>
The weight of a large number of fat persons is nicely modeled with a normal distribution with mean of 128kg and a standard deviation of 9 kg

1. The probability of fat persons with weight at most 110 kg is:

$$X \sim N(\ \mu\ \ , \sigma^2)$$
$$X \sim N(128, 9^2)$$

$$P(X \leq 110) = P\left(Z < \frac{110 - 128}{9}\right) = P(Z < -2) = 0.0228$$

2. The probability of fat persons with weight more than 149 kg is:

$$P(X > 149) = P\left(Z > \frac{149 - 128}{9}\right) = 1 - P(Z < 2.33) = 1 - 0.9901 = 0.0099$$

3. The weight x above which 86% of those persons will be:

$$P(X > x) = 0.86 \Rightarrow P(X < x) = 0.14 \Rightarrow P\left(Z < \frac{x - 128}{9}\right) = 0.14$$

by searching inside the table for 0.14, and transforming X to Z, we got:

$$\frac{x - 128}{9} = -1.08$$

$$x - 128 = -1.08 \times 9$$

$$x = (-1.08 \times 9) + 128$$

$$x = 118.28$$

4. The weight x below which 50% of those persons will be:

$P(X < x) = 0.5$, *by searching inside the table for 0.5, and transforming X to Z*

$$\frac{x - 128}{9} = 0 \Rightarrow x = 128$$

Question 6:

If the random variable X has a normal distribution with the mean μ and the variance σ^2, then $P(X < \mu + 2\sigma)$ equal to:

$$P(X < \mu + 2\sigma) = P\left(Z < \frac{(\mu + 2\sigma) - \mu}{\sigma}\right) = P(Z < 2) = 0.9772$$

Question 7:

If the random variable X has a normal distribution with the mean μ and the variance 1, and if then $P(X < 3) = 0.877$ then μ equal to

$$Given\ that\ \ \sigma = 1$$

$$P(X < 3) = 0.877 \Rightarrow P\left(Z < \frac{3-\mu}{1}\right) = 0.877$$

$$3 - \mu = 1.16 \Rightarrow \mu = 1.84$$

Question 8:

Suppose that the marks of students in a certain course are distributed according to a normal distribution with the mean 70 and the variance 25. If it is known that 33% of the student failed the exam, then the passing mark is:

$$X \sim N(70, 25)$$

$$P(X < x) = 0.33 \Rightarrow P\left(Z < \frac{x - 70}{5}\right) = 0.33$$

by searching inside the table for 0.33, and transforming X to Z, we got:

$$\frac{x - 70}{5} = -0.44 \Rightarrow x = 67.8$$

Question 9:

What k value corresponds to 17% of the area between the mean and the z value? $P(\mu < Z < k) = 0.17$

$$P(\mu < Z < k) = 0.17$$
$$P(Z < k) - P(Z < \mu) = 0.17$$
$$P(Z < k) - \quad 0.5 \quad = 0.17$$
$$P(Z < k) = 0.67$$
$$k = 0.44$$

Question 10:

A nurse supervisor has found that staff nurses complete a certain task in 10 minutes on average. If the times required to complete the task are approximately normally distributed with a standard deviation of 3 minutes, then:

1) The probability that a nurse will complete the task in less than 8 minutes is:

$$X \sim N(10, 3^2)$$

$$P(X < 8) = P\left(Z < \frac{8-10}{3}\right) = P(Z < -0.67) = 0.2514$$

2) The probability that a nurse will complete the task in more than 4 minutes is:

$$P(X > 4) = 1 - P\left(Z < \frac{4-10}{3}\right) = 1 - P(Z < -2) = 1 - 0.0228 = 0.9772$$

3) If eight nurses were assigned the task, the expected number of them who will complete it within 8 minutes is approximately equal to:

$$\begin{aligned} n \times P(0 < X < 8) &= 8 \times P\left(\frac{0-10}{3} < Z < \frac{8-10}{3}\right) \\ &= 8 \times P(-3.33 < Z < -0.67) \\ &= 8 \times [P(Z < -0.67) - P(Z < -3.33)] \\ &= 8 \times [0.2514 - 0.0004] = 2 \end{aligned}$$

4) If a certain nurse completes the task within k minutes with probability 0.6293; then k equals approximately:

$$\begin{aligned} &P(0 < X < k) = 0.6293 \\ &\Longrightarrow P\left(\frac{0-10}{3} < Z < \frac{k-10}{3}\right) = 0.6293 \\ &\Longrightarrow P\left(-3.33 < Z < \frac{k-10}{3}\right) = 0.6293 \\ &\Longrightarrow P\left(Z < \frac{k-10}{3}\right) - P(Z < -3.33) = 0.6293 \\ &\Longrightarrow P\left(Z < \frac{k-10}{3}\right) - 0.0004 = 0.6293 \\ &\Longrightarrow P\left(Z < \frac{k-10}{3}\right) = 0.6297 \\ &\Longrightarrow \frac{k-10}{3} = 0.33 \Longrightarrow k = 11 \end{aligned}$$

Question 11:

Given the normally distributed random variable X with mean 491 and standard deviation 119,

1. If P (X < k) = 0.9082, the value of k is equal to

A	649.27	B	390.58	C	128.90	D	132.65

2. If P (292< X <M) = 0.8607, the value of M is equal to

A	766	B	649	C	108	D	136

Question 12:

The IQ (Intelligent Quotient) of individuals admitted to a state school for the mentally retarded are approximately normally distributed with a mean of 60 and a standard deviation of 10, then:

1) The probability that an individual picked at random will have an IQ greater than 75 is:

A	0.9332	B	0.8691	C	0.7286	D	0.0668

2) The probability that an individual picked at random will have an IQ between 55 and 75 is:

A	0.3085	B	0.6915	C	0.6247	D	0.9332

3) If the probability that an individual picked at random will have an IQ less than k is 0.1587. Then the value of k

A	50	B	45	C	51	D	40

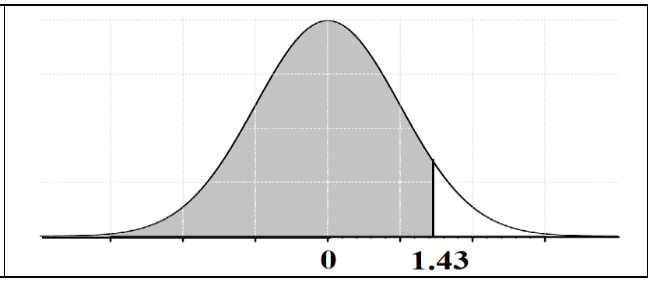
More Exercises:

Exercise 1:

Suppose that the random variable Z has a standard normal distribution

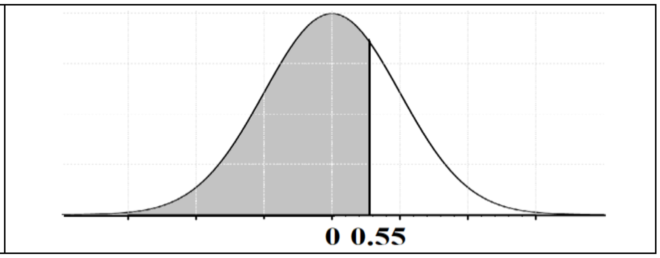
(a) Find the area to the left of Z =1.43.

$P(Z < 1.43) = 0.92364$



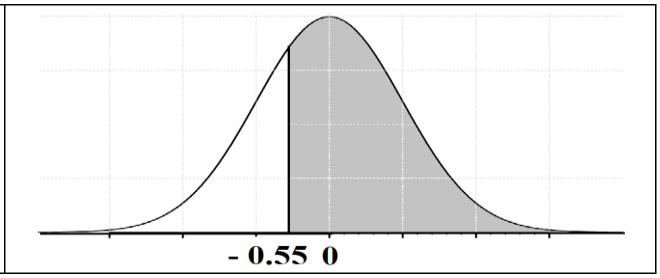
(b) Find P (Z < 0.55).

$P(Z < 0.55) = 0.70884$



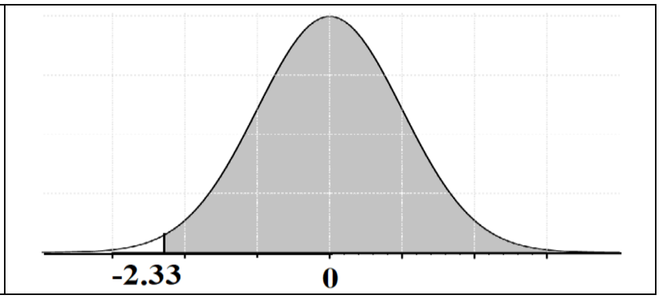
(c) Find P (Z > - 0.55).

$P(Z > -0.55)$
$=1-P(Z<-0.55)$
$=1-0.29116$
$= 0.70884$



(d) Find P (Z > - 2.33).

$P(Z > -2.33)$
$= 1-P(Z<-2.33) =1- 0.00990$
$= 0.9901$



(e) Find the area to the right of z = 2.

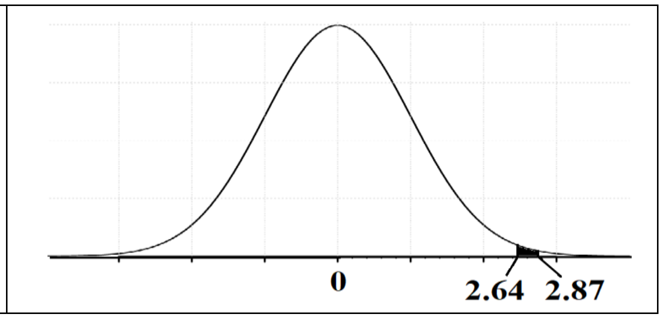
P(Z > 2.00) =1-P(Z<2.00)
= 1- 0.97725
= 0.02275

(f) Find the area under the curve between z = 0 and z = 1.43.

P(0 < Z <1.43)
= P(Z <1.43) – P(Z<0.00)
= 0.92364 - 0.5
= 0.42364

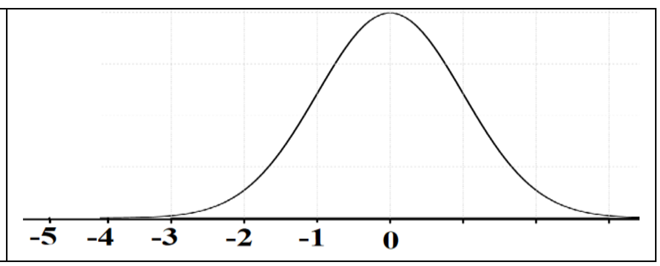
(g) Find the probability that Z will take a value between z = 2.64 and z = 2.87.

P(2.64 < Z <2.87)
= P(Z <2.87) – P(Z<2.64)
= 0.99795 -0.99585
= 0.0021



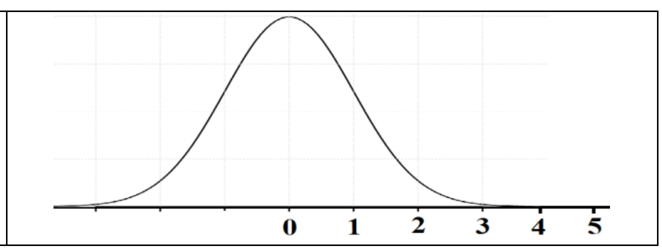
(h) Find P(Z < - 5).

P(Z < - 5) ≈ 0



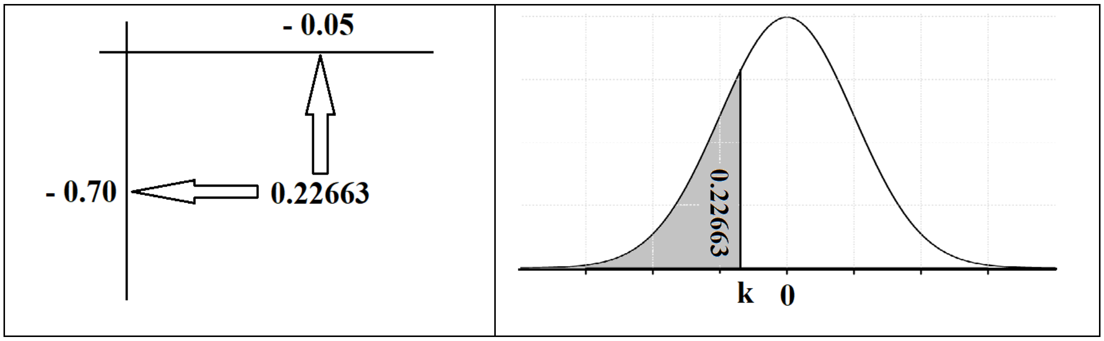
(i) Find P(Z > 5).

P(Z > 5) = 1-P(Z<5)
≈ 1 - 1
= 0



(j) If $P(Z \le k) = 0.22663$, then find the value of k.

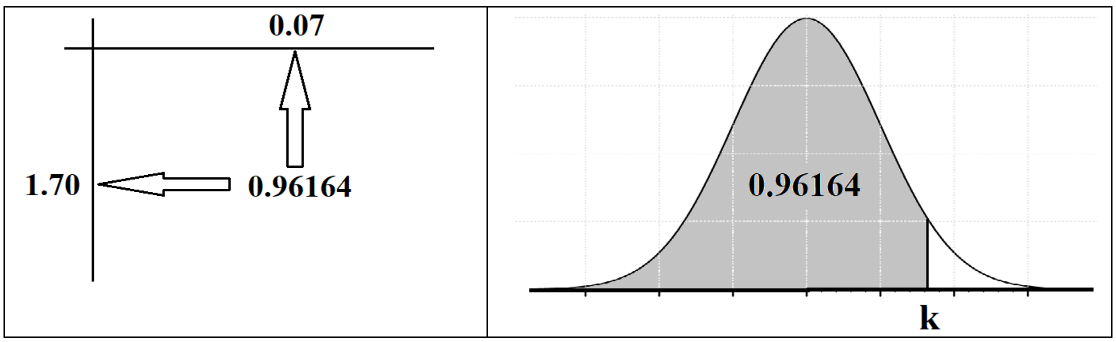
$k = -0.75$



(k) If $P(Z \ge k) = 0.03836$, then find the value of k.

$P(Z<k)=1- P(Z \ge k) =1-0.03836 = 0.96164$

$k = 1.77$



l) If $P(-2.67 < Z \le k) = 0.97179$, then find the value of k.

$0.97179 = P(-2.67 < Z \le k)$

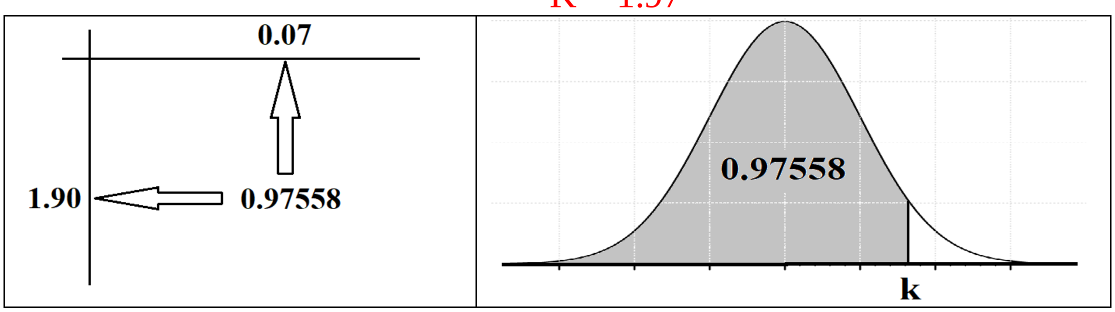
$= P(Z<K) – P(Z < -2.67)$

$P(Z<K) = 0.97179 + P(Z < -2.67)$

$P(Z<K) =0.97179 + 0.00379$

$= 0.97558$

$K = 1.97$



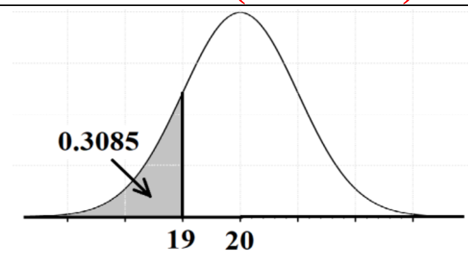
<u>Exercise 2</u>:

Suppose that the time for a person to be tested for corona virus (in minutes) has a normal distribution with mean $\mu = 20$ and variance $\sigma^2 = 4$.

(1) If we select a person at random, what is the probability that his examination period will be less than 19 minutes?

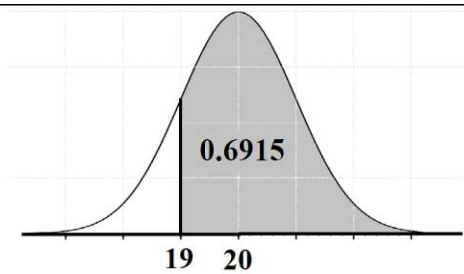
Let X = person's period of examination (in minutes)

$$P(X < 19) = P\left(\frac{X-\mu}{\sigma} < \frac{19-\mu}{\sigma}\right)$$
$$= P\left(Z < \frac{19-20}{2}\right)$$
$$= P(Z < -0.5)$$
$$= 0.3085$$



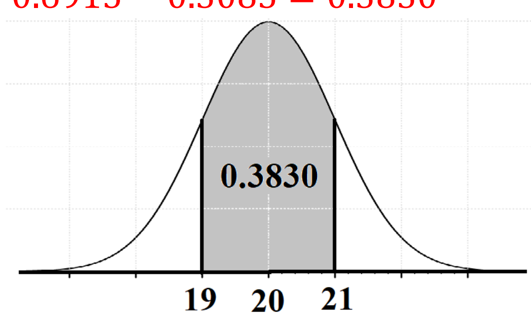
(2) If we select a person at random, what is the probability that his examination period will be more than 19 minutes?

$$P(X > 19) = 1 - P(X < 19)$$
$$= 1 - 0.3085$$
$$= 0.6915$$



(3) If we select a person at random, what is the probability that his examination period will be between 19 and 21 minutes?

$$P(19 < X < 21) = P(X < 21) - P(X < 19)$$
$$= P\left(\frac{X-\mu}{\sigma} < \frac{21-\mu}{\sigma}\right) - P\left(\frac{X-\mu}{\sigma} < \frac{19-\mu}{\sigma}\right)$$
$$= P\left(Z < \frac{21-20}{2}\right) - P\left(Z < \frac{19-20}{2}\right)$$
$$= P(Z < 0.5) - P(Z < -0.5)$$
$$= 0.6915 - 0.3085 = 0.3830$$



(4) What is the percentage of persons whose examination period are less than 19 minutes?

$$\% \ = \ P(X < 19) * 100\% = 0.3085 * 100\%$$
$$= 30.85\%$$

(5) If we select a sample of 2000 persons, how many persons would be expected to have examination periods that are less than 19 minutes?

$$\text{Expected number} = \ 2000 \ \times P(X < 19)$$
$$= 2000 \times 0.3085$$
$$= 617$$

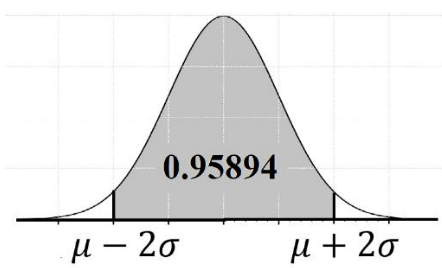
Exercise 3:

Suppose that we have a normal population with mean μ and standard deviation σ.

(1) Find the percentage of values which are between $\mu - 2\sigma$ and $\mu + 2\sigma$.

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

$$= \ P\left(\frac{X-\mu}{\sigma} < \frac{(\mu+2\sigma)-\mu}{\sigma}\right) - P\left(\frac{X-\mu}{\sigma} < \frac{(\mu-2\sigma)-\mu}{\sigma}\right)$$
$$= \ P\left(Z < \frac{(\mu+2\sigma)-\mu}{\sigma}\right) - P\left(Z < \frac{(\mu-2\sigma)-\mu}{\sigma}\right)$$
$$= \ P\left(Z < \frac{2\sigma}{\sigma}\right) - P\left(Z < \frac{-2\sigma}{\sigma}\right)$$
$$= \ P(Z < 2.00) - P(Z < -2.00)$$
$$= 0.97725 \ - 0.02275$$
$$= 0.9545$$



(2) Find the percentage of values which are between $\mu - \sigma$ and $\mu + \sigma$.

Dot it yourself

(3) Find the percentage of values which are between $\mu - 3\sigma$ and $\mu + 3\sigma$.

Dot it yourself

<u>Exercise 4:</u> (Read it yourself)

In a study of fingerprints, an important quantitative characteristic is the total ridge count for the 10 fingers of an individual. Suppose that the total ridge counts of individuals in a certain population are approximately normally distributed with a mean of 140 and a standard deviation of 50. Then:

(1) The probability that an individual picked at random from this population will have a ridge count of 200 or more is:

$$\begin{aligned} P(X > 200) &= 1 - P(X < 200) \\ &= 1 - P\left(Z < \tfrac{200-\mu}{\sigma}\right) \\ &= 1 - P\left(Z < \tfrac{200-140}{50}\right) \\ &= 1 - P(Z < 1.2) \\ &= 1 - 0.88493 = 0.11507. \end{aligned}$$

(2) The probability that an individual picked at random from this population will have a ridge count of less than 100 is:

$$\begin{aligned} P(X < 100) &= P\left(Z < \frac{100-\mu}{\sigma}\right) \\ &= P\left(Z < \tfrac{100-140}{50}\right) \\ &= P(Z < -0.80) = 0.18673 \end{aligned}$$

(3) The probability that an individual picked at random from this population will have a ridge count between 100 and 200 is:

$$\begin{aligned} P(100 < X < 200) &= P(X < 200) - P(X < 100) \\ &= P(X < 200) - P(X < 100) \\ &= P\left(Z < \tfrac{200-140}{50}\right) - \left(Z < \tfrac{100-140}{50}\right) \\ &= P(Z < 1.20) - (Z < -0.80) \\ &= 0.88493 - 0.18673 = 0.6982 \end{aligned}$$

(4) The percentage of individuals whose ridge counts are between 100 and 200 is:

$$\begin{aligned} P(100 < X < 200) * 100\% &= 0.6982 * 100\% \\ &= 69.82\% \end{aligned}$$

(4) If we select a sample of 5,000 individuals from this population, how many individuals would be expected to have ridge counts that are between 100 and 200?

$$\begin{aligned} \text{Expected number} &= 5000 \times P(100 < X < 200) \\ &= 5000 \times 0.6982 = 3491 \end{aligned}$$

Chapter 5 Sampling Distribution

Sampling Distribution

Single Mean	Two Means
$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$ $E(\bar{X}) = \bar{X} = \mu$ $Var(\bar{X}) = \frac{\sigma^2}{n}$	$$\bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2 , \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$ $E(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2$ $Var(\bar{X}_1 - \bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$
Single Proportion	**Two Proportions**
For large sample size ($n \geq 30,\ np > 5,\ nq > 5$) $$\hat{p} \sim N\left(p, \frac{pq}{n}\right)$$ $E(\hat{p}) = p$ $Var(\hat{p}) = \frac{pq}{n}$	For large sample size ($n_1 \geq 30,\ n_1p_1 > 5,\ n_1q_1 > 5$) ($n_2 \geq 30,\ n_2p_2 > 5,\ n_2q_2 > 5$) $$\hat{p}_1 - \hat{p}_2 \sim N\left(p_1 - p_2, \frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}\right)$$ $E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2$ $Var(\hat{p}_1 - \hat{p}_2) = \frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}$

	population normal or not normal n large ($n \geq 30$)		population normal n small ($n < 30$)	
	σ known	σ unknown	σ known	σ unknown
Sampling Distribution	$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$	$Z = \frac{\bar{X} - \mu}{s/\sqrt{n}}$	$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$	$T = \frac{\bar{X} - \mu}{s/\sqrt{n}}$

Question 1:

The average life of a certain battery is 5 years, with a standard deviation of 1 year. Assume that the live of the battery approximately follows a normal distribution.

1. The sample mean $\bar{X}$ of a random sample of 5 batteries selected from this product has mean $E(\bar{X}) = \mu_{\bar{X}}$.

$$\boxed{\mu = 5 \;\; ; \;\; \sigma = 1 \;\; ; \;\; n = 5}$$
$$E(\bar{X}) = \mu = 5$$

2. The variance $Var(\bar{X}) = \sigma_{\bar{X}}^2$ of the sample mean $\bar{X}$ of a random sample of 5 batteries selected from this product is equal to:

$$Var(\bar{X}) = \frac{\sigma^2}{n} = \frac{1}{5} = 0.2$$

3. The probability that the average life of a random sample of size 16 of such batteries will be between 4.5 and 5.4.

$$n = 16 \rightarrow \frac{\sigma}{\sqrt{n}} = \frac{1}{4}$$

$$P(4.5 < \bar{X} < 5.4) = P\left(\frac{4.5-\mu}{\frac{\sigma}{\sqrt{n}}} < Z < \frac{5.4-\mu}{\frac{\sigma}{\sqrt{n}}}\right)$$
$$= P\left(\frac{4.5-5}{\frac{1}{4}} < Z < \frac{5.4-5}{\frac{1}{4}}\right) = P(-2 < Z < 1.6)$$
$$= P(Z < 1.6) - P(Z < -2)$$
$$= 0.9452 - 0.0228 = 0.9224$$

4. The probability that the average life of a random sample of size 16 of such batteries will be less than 5.5 years is:

$$P(\bar{X} < 5.5) = P\left(Z < \frac{5.5-\mu}{\frac{\sigma}{\sqrt{n}}}\right) = P\left(Z < \frac{5.5-5}{1/4}\right) = P(Z < 2) = 0.9772$$

5. The probability that the average life of a random sample of size 16 of such batteries will be more than 4.75 years is:

$$P(\bar{X} > 4.75) = P\left(Z > \frac{4.75-\mu}{\frac{\sigma}{\sqrt{n}}}\right)$$
$$= P\left(Z > \frac{4.75-5}{\frac{1}{4}}\right) = P(Z > -1)$$
$$= 1 - P(Z < -1) = 1 - 0.1587 = 0.841$$

6. If $P(\bar{X} > a) = 0.1492$ where $\bar{X}$ represent the sample mean for a random sample of size 9 of such batteries, then the numerical value of a is:

$P(\bar{X} > a) = 0.1492 \;\; ; \;\; n = 9$

$$P\left(Z > \frac{a-\mu}{\frac{\sigma}{\sqrt{n}}}\right) = 0.1492$$

$$\Rightarrow P\left(Z < \frac{a-5}{\frac{1}{3}}\right) = 1 - 0.1492$$

$$\Rightarrow P\left(Z < \frac{a-5}{\frac{1}{3}}\right) = 0.8508$$

$$\frac{a-5}{\frac{1}{3}} = 1.04 \Longrightarrow a = 5.347$$

<u>Question 2:</u>

Suppose that you take a random sample of size $n = 64$ from a distribution with mean $\mu = 55$ and standard deviation $\sigma = 10$. Let $\bar{X} = \frac{1}{n}\sum x$ be the sample mean.

1. What is the approximate sampling distribution of $\bar{X}$.

$$\boxed{\mu = 55 \;\; ; \;\; \sigma = 10 \;\; ; \;\; n = 64}$$

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) = \bar{X} \sim N\left(55, \frac{100}{64}\right)$$

2. What is the mean of $\bar{X}$?

$$E(\bar{X}) = \mu = 55$$

3. What is the standard error (standard deviation) of $\bar{X}$?

$$S.D(\bar{X}) = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{64}} = \frac{10}{8}$$

4. Find the probability that the sample mean $\bar{X}$ exceeds 52.

$$P(\bar{X} > 52) = P\left(Z > \frac{52-55}{\frac{10}{8}}\right) = P(Z > -2.4)$$

$$= 1 - P(Z < -2.4)$$

$$= 1 - 0.0082 = 0.9918$$

Question 3:

Suppose that the hemoglobin levels (in g/dl) of healthy Saudi females are approximately normally distributed with mean of 13.5 and a standard deviation of 0.7. If 15 healthy adult Saudi female is randomly chosen, then:

1. The mean of $\overline{X}$ ($E(\overline{X})$ or $\mu_{\overline{X}}$)

A	0.7	B	13.5	C	15	D	3.48

2. The standard error of $\overline{X}$ ($\sigma_{\overline{X}}$)

A	0.181	B	0.0327	C	0.7	D	13.5

3. $P(\overline{X} < 14) =$

A	0.99720	B	0.99440	C	0.76115	D	0.9971

4. $P(\overline{X} > 13.5) =$

A	0.99	B	0.50	C	0.761	D	0.622

5. $P(13 < \overline{X} < 14) =$

A	0.9972	B	0.9944	C	0.7615	D	0.5231

Question 4:

If the uric acid value in normal adult males is approximately normally distributed with a mean and standard derivation of 5.7 and 1 mg percent, respectively, find the probability that a sample of size 9 will yield a mean

1. Greater than 6 is:

A	0.2109	B	0.1841	C	0.8001	D	0.8159

2. At most 5.2 is:

A	0.6915	B	0.9331	C	0.8251	D	0.0668

3. Between 5 and 6 is:

A	0.1662	B	0.7981	C	0.8791	D	0.9812

Question 5:

Medical research has concluded that people experience a common cold roughly two times per year. Assume that the time between colds is normally distributed with a mean 165 days and a standard deviation of 45 days. Consider the sampling distribution of the sample mean based on samples of size 36 drown from the population:

1. The mean of sampling distribution $\bar{X}$ is:

A	210	B	36	C	45	D	165

2. The distribution if the mean of $\bar{X}$ is:

A	$N(165{,}2025)$	B	$N(165{,}45)$	C	$T, with\ df = 30$	D	$N(165{,}7.5)$

3. $\mathrm{P}(\bar{\mathrm{X}} > 178) =$

A	0.0415	B	0.615	C	0.958	D	0.386

Sampling Distribution: Two Means:

$$* \bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2\ , \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

$$* E(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2 \qquad * Var(\bar{X}_1 - \bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

Question 6:

A random sample of size $n_1 = 36$ is taken from normal population with a mean $\mu_1 = 70$ and a standard deviation $\sigma_1 = 4$. A second independent random sample of size $n_2 = 49$ is taken from a normal population with a mean $\mu_2 = 85$ and a standard deviation $\sigma_2 = 5$. Let $\bar{X}_1$ and $\bar{X}_2$ be the average of the first and second sample, respectively.

1. Find $E(\bar{X}_1 - \bar{X}_2)$ and $Var(\bar{X}_1 - \bar{X}_2)$.

$$n_1 = 36\,, \mu_1 = 70\,, \sigma_1 = 4$$
$$n_2 = 49\,, \mu_2 = 85\,, \sigma_2 = 5$$

$$E(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2 = 70 - 85 = -15$$

$$Var(\bar{X}_1 - \bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{16}{36} + \frac{25}{49} = 0.955$$

2. Find $P(\bar{X}_1 - \bar{X}_2 > -16)$.

$$P(\bar{X}_1 - \bar{X}_2 > -16) = P\left(Z > \frac{-16-(-15)}{\sqrt{0.955}}\right) = 1 - P\left(Z < \frac{-16-(-15)}{\sqrt{0.955}}\right)$$

$$= 1 - P(Z < -1.02) = 0.8461$$

Question 7:

A random sample of size 25 is taken from a normal population (1st population) having a mean of 100 and a standard of 6. A second random sample of size 36 is taken from a different normal population (2nd population) having a mean of 97 and a standard deviation of 5. Assume that these two samples are independent.

1. The probability that the sample mean of the first sample will exceed the sample mean of the second sample by at least 6 is:

$$n_1 = 25\,,\ \mu_1 = 100\,, \sigma_1 = 6$$
$$n_2 = 36\,,\ \mu_2 = 97\ \ , \sigma_2 = 5$$

$$E(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2 = 100 - 97 = 3 \qquad Var(\bar{X}_1 - \bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{36}{25} + \frac{25}{36} = 2.134$$

$$P(\bar{X}_1 > \bar{X}_2 + 6) = P(\bar{X}_1 - \bar{X}_2 > 6)$$

$$= P\left(Z > \frac{6-(3)}{\sqrt{2.134}}\right) = P(Z > 2.05)$$
$$= 1 - P(Z < 2.05)$$
$$= 1 - 0.9798 = 0.0202$$

2. The probability that the difference between the two-sample means will be less than 2 is:

$$P(\bar{X}_1 - \bar{X}_2 < 2) = P\left(Z < \frac{2-(3)}{\sqrt{2.134}}\right)$$

$$= P(Z < -0.68) = 0.2483$$

Question 8:

Given two normally distributed population with equal means and variances $\sigma_1^2 = 100$, $\sigma_2^2 = 350$. Two random samples of size $n_1 = 40$, $n_2 = 35$ are drown and sample means $\bar{X}_1$ and $\bar{X}_2$ are calculated, respectively, then

$\boxed{\mu_1 - \mu_2 = 0}$

1. $P(\bar{X}_1 - \bar{X}_2 > 12)$ is

A	0.1499	B	0.8501	C	0.9997	D	0.0003

2. $P(5 < \bar{X}_1 - \bar{X}_2 < 12)$ is

A	0.0789	B	0.9217	C	0.8002	D	None of these

Sampling Distribution: Single Proportion:

For large sample size ($n \geq 30,\ np > 5,\ nq > 5$)

$$* \hat{p} \sim N\left(p, \frac{pq}{n}\right)$$

$$* E(\hat{p}) = p \qquad * Var(\hat{p}) = \frac{pq}{n}$$

Question 9:

Suppose that 10% of the students in a certain university smoke cigarette. A random sample of 30 student is taken from this university. Let $\hat{p}$ be the proportion of smokers in the sample.

1. Find $E(\hat{p}) = \mu_{\hat{p}}$ the mean of $\hat{p}$.

$$\boxed{p = 0.1 \ ; \ n = 30 \ ; \ q = 1 - p = 0.9}$$

$$E(\hat{p}) = p = 0.1$$

2. Find $Var(\hat{p}) = \sigma^2_{\hat{p}}$ the variance of $\hat{p}$.

$$Var(\hat{p}) = \frac{pq}{n} = \frac{0.1 \times 0.9}{30} = 0.003$$

3. Find an approximate distribution of $\hat{p}$

$$\hat{p} \sim N(0.1, 0.003)$$

4. Find $P(\hat{p} > 0.25)$.

$$P(\hat{p} > 0.25) = P\left(Z > \frac{0.25 - 0.1}{\sqrt{0.003}}\right) = P(Z > 2.74)$$

$$= 1 - P(Z < 2.74) = 1 - 0.99693 = 0.00307$$

Question 10:
Suppose that 15% of the patients visiting a certain clinic are females. If A random sample of 35 patients was selected, $\hat{p}$ represent the proportion of females in the sample. then find:

1. The expected value of $(\hat{p})$ is:

A	0.35	B	0.85	C	0.15	D	0.80

2. The standard deviation of $(\hat{p})$ is:

A	0.3214	B	0.0036	C	0.1275	D	0.0604

3. The approximate sampling distribution of $(\hat{p})$ is:

A	N(0.15,0.0604)	B	Binomial(0.15,35)	C	$N(0.15, 0.0604^2)$	D	$Binomial(0.15, 35^2)$

4. The $P(\hat{p} > 0.15)$ is:

A	0.35478	B	0.5	C	0.96242	D	0.46588

Question 11:
In a study, it was found that 31% of the adult population in a certain city has a diabetic disease. 100 people are randomly sampled from the population. Then

1. The mean for the sample proportion $\left(E(\hat{p}) \text{ or } \mu_{\hat{p}}\right)$ is:

A	0.40	B	0.31	C	0.69	D	0.10

2. $P(\hat{p} > 0.40) =$

A	0.02619	B	0.02442	C	0.0256	D	0.7054

Sampling Distribution: Two Proportions:

For large sample size $(n_1 \geq 30,\ n_1p_1 > 5,\ n_1q_1 > 5)$
$(n_2 \geq 30,\ n_2p_2 > 5,\ n_2q_2 > 5)$

$$* \hat{p}_1 - \hat{p}_2 \sim N\left(p_1 - p_2, \frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}\right)$$

$$* E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2 \qquad * Var(\hat{p}_1 - \hat{p}_2) = \frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}$$

Question 12:

Suppose that 25% of the male student and 20% of the female student in certain university smoke cigarettes. A random sample of 35 male students is taken. Another random sample of 30 female student is independently taken from this university. Let $\hat{p}_1$ and $\hat{p}_2$ be the proportions of smokers in the two sample, respectively.

1. Find $E(\hat{p}_1 - \hat{p}_2) = \mu_{\hat{p}_1 - \hat{p}_2}$, the mean of $\hat{p}_1 - \hat{p}_2$.

$$p_1 = 0.25 \ ; \quad n_1 = 35$$
$$p_2 = 0.20 \ ; \quad n_2 = 30$$

$$E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2 = 0.25 - 0.20 = 0.05$$

2. Find $Var(\hat{p}_1 - \hat{p}_2) = \sigma^2_{\hat{p}_1 - \hat{p}_2}$, the variance of $\hat{p}_1 - \hat{p}_2$.

$$Var(\hat{p}_1 - \hat{p}_2) = \frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2} = \frac{0.25 \times 0.75}{35} + \frac{0.2 \times 0.8}{30} = 0.01069$$

3. Find an approximate distribution of $\hat{p}_1 - \hat{p}_2$.

$$\hat{p}_1 - \hat{p}_2 \sim N(0.05, 0.01069)$$

4. Find $P(0.1 < \hat{p}_1 - \hat{p}_2 < 0.2)$

$$P(0.1 < \hat{p}_1 - \hat{p}_2 < 0.2) = \left(\frac{0.1 - 0.05}{\sqrt{0.01069}} < Z < \frac{0.2 - 0.05}{\sqrt{0.01069}}\right)$$

$$= \quad (0.48 < Z < 1.45)$$
$$= P(Z < 1.45) - P(Z < 0.48)$$
$$= 0.92647 - 0.68439 = 0.24208$$

Question 13:

Suppose that 7 % of the pieces from a production process A are defective while that proportion of defective for another production process B is 5 %. A random sample of size 400 pieces is taken from the production process A while the sample size taken from the production process B is 300 pieces. If $\hat{p}_1$ and $\hat{p}_2$ be the proportions of defective pieces in the two samples, respectively, then:

1. The sampling distribution of $(\hat{p}_1 - \hat{p}_2)$ is:

A	N(0,1)	B	Normal	C	T	D	Unknown

2. The value of the standard error of the difference $(\hat{p}_1 - \hat{p}_2)$ is:

A	0.02	B	0.10	C	0	D	0.22

Question 13:

In a study to make an inference between the proportion of houses heated by gas in city A and city B, the following information was collected:

	Proportion of houses heated by gas	Sample size
City A	43%	90
City B	51%	150

Suppose p_A proportion of city A houses which are heated by gas, p_B proportion of city B houses which are heated by gas. The two sample are independent.

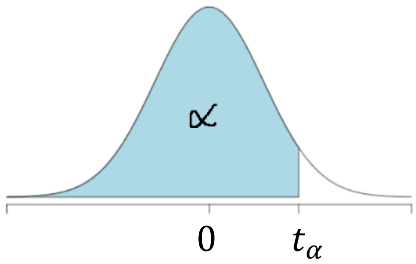
1. The sampling distribution for the sample proportion of city B which are heated by gas is:

A	$\hat{p}_B \sim N\left(p_B, \frac{p_B q_B}{n_B}\right)$	B	$\hat{p}_B \sim N\left(\hat{p}_B, \frac{\hat{p}_B \hat{q}_B}{n_B}\right)$	C	$\hat{p}_B \sim N(\hat{p}_B, \hat{p}_B \hat{q}_B)$	D	$\hat{p}_B \sim N(p_B, p_B q_B)$

2. The sampling distribution of $\hat{p}_A$ is (approximately) normal if:

A	$n_A \geq 30$ $n_A p_A > 5$	B	$n_A \geq 30$ $n_A p_A > 5$ $n_A q_A > 5$	C	$n_A p_A > 5$	D	$\frac{p_A}{n_A} > 5$

The student (t) Distribution:



يجب ان تكون اشارة الاحتمال أقل من (<) قبل البحث في جدول (t) :

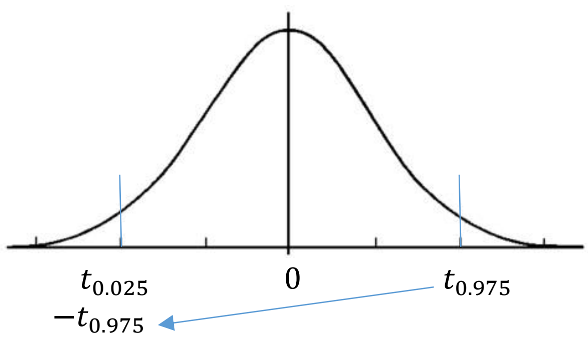
$$t_{\alpha,v} \Longrightarrow P(T < t_{\alpha,v})$$
$$\Longrightarrow P(T < -t_{1-\alpha,v}) \qquad ; v = n-1$$

- $If\ P(T < t_{0.99,22}) \Longrightarrow t_{0.99,22} = 2.508$

- $If\ P(T > t_{0.975,18}) \Longrightarrow P(T < t_{0.025,18})$

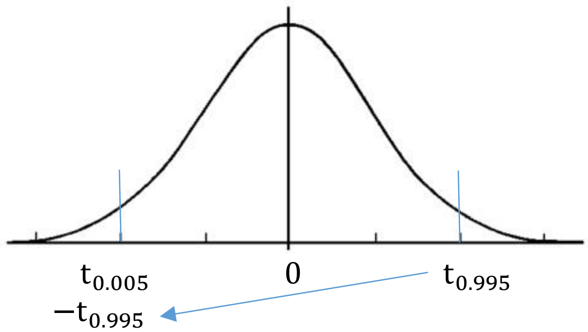
$$\Longrightarrow P(T < -t_{0.975,18})$$

$$\Longrightarrow -t_{0.975,18} = -2.101$$



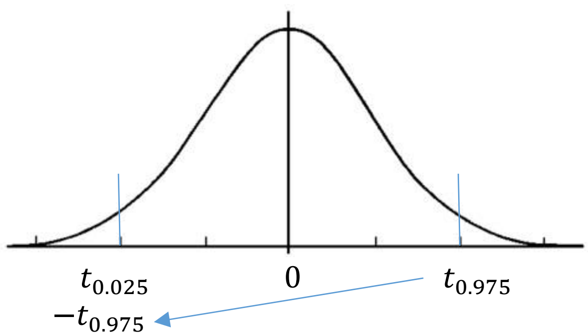
- If $P(T > t_{\alpha,v})$ where $v = 24, \alpha = 0.995$

$$\Rightarrow P(T > t_{0.995,24})$$
$$\Rightarrow P(T < t_{0.005,24})$$
$$\Rightarrow P(T < -t_{0.995,24})$$
$$\Rightarrow -t_{0.995,24} = -2.797$$



- If $P(T > t_{\alpha,v})$ where $v = 7, \alpha = 0.975$

$$\Rightarrow P(T > t_{0.975\ 7})$$
$$\Rightarrow P(T < t_{0.025\ 7})$$
$$\Rightarrow P(T < -t_{0.975,7})$$
$$\Rightarrow -t_{0.975,7} = -2.365$$



Question 1:
Let T follow the t distribution with 9 degrees of freedom, then
The probability $(T < 1.833)$ equal to:

بما ان الاشارة اقل من (<) اذن ننظر للجدول

ν=df	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.99}$	$t_{0.995}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055

$\alpha = 0.95$

- $The\ probabitity\ P(T < -1.833)\ equal\ to$:

ν=df	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.99}$	$t_{0.995}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055

$\alpha = 1 - 0.95 = 0.05$

Question 2:
Let T follow the t distribution with 9 degrees of freedom, then
The t-value that leaves an area of 0.10 to the right is:

$$P(T > t_{0.10,9})$$
$$P(T < t_{0.90,9}) \implies t_{0.90,9} = 1.383$$

ν=df	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.99}$	$t_{0.995}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055

Question 3:
Given the t-distribution with 12 degrees of freedom, then
The t-value that leaves an area of 0.025 to the left is:

$$P(T < t_{0.025,12})$$
$$P(T < -t_{0.975,12})$$

ν=df	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.99}$	$t_{0.995}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055

$$-t_{0.975,12} = -2.179$$

Question 4:

Consider the student t distribution:

Find the t-value with $\boldsymbol{n = 17}$ the leaves an area of 0.01 to the left:

$\text{df} = \ \text{n} - 1$

$= 17 - 1 = 16$

$P(T < \ \ t_{0.01,16})$

$P(T < -t_{0.99,16})$

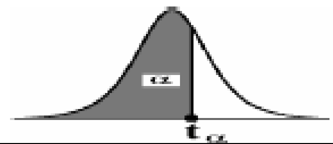
v=df	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.99}$	$t_{0.995}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861

$t_{0.99,16} = \ \ 2.583$

$-t_{0.99,16} = -2.583$

A	-2.58	B	2.567	C	2.58	D	-2.567

Critical Values of the t-distribution (t_α)



v=df	$t_{0.90}$	$t_{0.95}$	$t_{0.975}$	$t_{0.99}$	$t_{0.995}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
35	1.3062	1.6896	2.0301	2.4377	2.7238
40	1.3030	1.6840	2.0210	2.4230	2.7040
45	1.3006	1.6794	2.0141	2.4121	2.6896
50	1.2987	1.6759	2.0086	2.4033	2.6778
60	1.2958	1.6706	2.0003	2.3901	2.6603
70	1.2938	1.6669	1.9944	2.3808	2.6479
80	1.2922	1.6641	1.9901	2.3739	2.6387

- **Question from previous midterms and finals:**

<u>Question:</u>
Given two normally distributed populations with a mean $\mu_1 = 10$ and $\mu_2 = 20$, and variances of $\sigma_1^2 = 100$ and $\sigma_2^2 = 80$. If two samples are taken from the populations of size $n_1 = 25$ and $n_1 = 16$ are taken from the populations.
Let $\overline{X}_1$ and $\overline{X}_2$ be the average of the first and second sample, respectively.

$$n_1 = 25\,, \mu_1 = 10\,, \sigma_1^2 = 100$$
$$n_2 = 16\,, \mu_2 = 20\,, \sigma_2^2 = 80$$

1. Find the sampling distribution for $\overline{\mathbf{X}}_1$.

$$\bar{X}_1 \sim N\left(\mu_1, \frac{\sigma_1^2}{n_1}\right)$$

$$\bar{X}_1 \sim N\left(10, \frac{100}{25}\right)$$

$$\bar{X}_1 \sim N(10, \ 4)$$

2. Find the sampling distribution for $(\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2)$.

$$\bar{X}_1 - \bar{X}_2 \sim N\left(\mu_1 - \mu_2\ , \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

$$\bar{X}_1 - \bar{X}_2 \sim N\left(10 - 20\ , \frac{100}{25} + \frac{80}{16}\right)$$

$$\bar{X}_1 - \bar{X}_2 \sim N(-10 \quad , \ 4 \quad + 5)$$

$$\bar{X}_1 - \bar{X}_2 \sim N(-10, 9)$$