

Exercise # 1

Q1: For each of the following variables indicate whether it is quantitative or qualitative variable:

- (a) The blood type of some patient in the hospital
 - (b) Blood pressure level of a patient.
 - (c) Weights of babies born in a hospital during a year.
 - (d) Gender of babies born in a hospital during a year.
 - (e) The distance between the hospital to the house
 - (f) Under-arm temperature of day-old infants born in a hospital.
-

Q2: For each of the following situations, answer questions (a) through (d):

- (a) What is the population?
- (b) What is the sample in the study?
- (c) What is the variable of interest?
- (d) What is the type of the variable?

Situation A: A study of 300 households in a small southern town revealed that if she has school-age child present.

- a-
- b-
- c-
- d-

Situation B: A study of 250 patients admitted to a hospital during the past year revealed that, Distance the patient live away from the hospital .

e-

f-

g-

h-

Q3:Choose the right answer:

1-The variable is a

- a. subset of the population.
- b. parameter of the population.
- c. relative frequency.
- d. characteristic of the population to be measured.
- e. class interval.

2-Which of the following is an example of discrete variable

- a. the number of students taking statistics in this term at KSU.
- b. the time to exercise daily.
- c. whether or not someone has a disease.
- d. height of certain buildings.
- e. Level of education.

3-Which of the following is not an example of discrete variable

- a. the number of students at the class of statistics.
- b. the number of times a child cry in a certain street.
- c. the time to run a certain distance.
- d. the number of buildings in a certain street.
- e. number of educated persons in a family.

4-Which of the following is an example of qualitative variable

- a. the blood pressure.
- b. the number of times a child brush his/her teeth.
- c. whether or not someone fail in an exam.
- d. Weight of babies at birth.
- e. the time to run a certain distance.

5-The continuous variable is a

- a. variable with a specific number of values.
- b. variable which can't be measured.
- c. variable takes on values within intervals.
- d. variable with no mode.
- e. qualitative variable.

6- which of the following is an example of continuous variable

- a. The number of visitors of the clinic yesterday.
- b. The time to finish the exam.
- c. The number of patients suffering from certain disease.
- d. Whether or not the answer is true.

7- The discrete variable is

- a-qualitative variable.
- b-variable takes on values within interval.
- c-variable with a specific number of values.

8-Which of the following is an example of nominal variable :

- a-age of visitors of a clinic.
- b-The time to finish the exam.
- c-Whether or not a person is infected by influenza.
- d-Weight for a sample of girls .

9-The nominal variable is a

- a-A variable with a specific number of values
- b-Qualitative variable that can't be ordered.
- c-variable takes on values within interval.
- d-Quantitative variable .

10-Which of the following is an example of nominal variable :

- a-The number of persons who are injured in accident.
- b-The time to finish the exam.
- c-Whether or not the medicine is effective.
- d-Socio-economic level.

11-The ordinal variable is :

- a-variable with a specific number of values.
- b-variable takes on values within interval.
- c-Qualitative variable that can be ordered.

Q1. A study was conducted in which they measured incidental intracranial aneurysms (IIAs) in 159 patients. The researchers examined complications and concluded that IIAs can be safely treated without causing mortality and with a lower complications rate than previously reported.

The following table represent the sizes (in millimeters) of the 159 IIAs in the sample:

IIAs size	frequency	Cumulative Frequency	Relative frequency	Cumulative Relative Frequency	Percentage frequency
0 - 4	29		0.182		
5 - 9	87		-----		
10 - 14	----		0.164		
15 - 19	10	-----	0.063		
20 - 24	4		0.025		-----
25 - 29	1		0.006		
30 - 34	2		0.013	----	
Total	-----				

Complete the table , then answer the following questions

1. The variable is , the type of variable is ----- ..
2. The number of patient with IIAs size between 10 – 14 is
3. The number of patient with IIAs size less than or equal 19 is
4. The relative frequency patient with IIAs size between 5 - 9 is
5. The proportion of patient with IIAs size less than 15 is
6. The percentage of patient with IIAs size between 15 – 29 is.....
7. The true class interval of (20 – 24) is
8. Width is
9. Maximum value is

Q2: The following table shows the hours 45 hospitals patients slept following the administration of a certain anesthetic .

True Class interval	Frequency	Midpoint
0.5 – 5.5	21	
5.5 – 10.5	16	-----
-----	-----	
15.5 – 20.5	2	
Total	45	

Answer the following questions:

1. The variable is

The type of variable is

2. The sample size is

3. The midpoint for the interval 5.5 – 10.5 is

4. The number of patients spend less than or equal 15.5 hour is

.....

5. The relative frequency of patients spend between 0.5 -10.5 hour is

.....

6. The class interval for the true class interval (5.5 – 10.5) is

..... (where $\frac{d}{2}=0.5$)

7. The percentage of patients spend more than 10.5 hour is

8. Width is

Q3: For a sample of patients, we obtain the following graph for approximate hours spend without pain after certain surgery .



Answer the following questions:

1. The type of the graph is
2. The variable is The type of the variable is
3. The sample size is
4. The number of patients spend a round 2 hours without pain is
5. The percent of patients spend 3.5 hours or more without pain is
6. The number of patients stayed the longest time without pain is
7. The lowest number of hours spent without pain is
8. Width is

Extra question 1:

Suppose two samples of human males yield the following data (which is more variation)

	Sample 1 25 year	Sample 2 11 year
Mean weight	135 pound	60 pound
Standard deviation	10 pound	10 pound

Sample 2 has variation than sample 1

Question 2:

The following values are calculated in respect of heights and weights for sample of students, can we say that the weights show greater variation than the heights.

	Sample 1 height	Sample 2 weight
Mean	162.6 cm	52.36 kg
variance	127.69 cm ²	23.14 kg ²

Since CV_2 than CV_1 , therefore we can say the weights show variability than heights.

شرح اضافي لأمثلة الاحتمالات (Probability)

Example : (Page 39)

Find: (سؤال اضافي)

$$\begin{aligned}P(A_3^C \cap B_2) &= P(A_1 \cap B_2) + P(A_2 \cap B_2) + P(A_4 \cap B_2) \\&= 9/339 + 30/339 + 0/339 \\&= 39/339 \\&= 0.11504\end{aligned}$$

$$\begin{aligned}P(A_2 \cap B_1^C) &= P(A_2 \cap B_2) + P(A_2 \cap B_3) \\&= 30/339 + 49/339 \\&= 79/339 \\&= 0.2330\end{aligned}$$

Example : (Page 44)

$P(A) = 0.9$, $P(B/A) = 0.8$ find $P(A \cap B) = ?$

Solution :

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

$$0.8 = \frac{P(A \cap B)}{0.9}$$

$$\begin{aligned}P(A \cap B) &= 0.8 \times 0.9 \\&= 0.72\end{aligned}$$

Bayes Theorm

The result of the test

	Has the disease (D)	Dose not have the disease (\bar{D})	Total
Positive (T)	Correct decision Sensitivity $P(T/D) = \frac{n(T \cap D)}{n(D)}$	False decision false positive $P(T/\bar{D}) = \frac{n(T \cap \bar{D})}{n(\bar{D})}$	n(T)
Negative (\bar{T})	False decision false negative $P(\bar{T}/D) = \frac{n(\bar{T} \cap D)}{n(D)}$	Correct decision Specificity $P(\bar{T}/\bar{D}) = \frac{n(\bar{T} \cap \bar{D})}{n(\bar{D})}$	n(\bar{T})
Total	n(D)	n(\bar{D})	n(Ω)

Predictive value Positive :

$$\begin{aligned} P(D/T) &= \frac{P(T/D) * P(D)}{P(T/D) * P(D) + P(T/\bar{D}) * P(\bar{D})} \\ &= \frac{P(T/D) * P(D)}{P(T/D) * P(D) + P(T/\bar{D}) * P(\bar{D})} = \\ &= \frac{Sensitivity * P(D)}{Sensitivity * P(D) + (1 - Specificity) * P(\bar{D})} \end{aligned}$$

Predictive value Negative :

$$\begin{aligned} P(\bar{D}/\bar{T}) &= \frac{P(\bar{T}/\bar{D}) * P(\bar{D})}{P(\bar{T}/\bar{D}) * P(\bar{D}) + P(\bar{T}/D) * P(D)} \\ &= \frac{P(\bar{T}/\bar{D}) * P(\bar{D})}{P(\bar{T}/\bar{D}) * P(\bar{D}) + P(\bar{T}/D) * P(D)} = \\ &= \frac{Specificity * P(\bar{D})}{Specificity * P(\bar{D}) + (1 - Sensitivity) * P(D)} \end{aligned}$$

Cumulative: (Example Page 60)

x	P(X≤x) : cumulative	P(X= x)
0	0.05	0.05
1	0.30	$0.30 - 0.05 = 0.25$
2	0.75	$0.75 - 0.30 = 0.45$
3	1	$1 - 0.75 = 0.25$
total	xxxxxxx	Total=1

x	P(X= x)
0	0.05
1	0.25
2	0.45
3	0.25
total	1

Extra example:

x	P(X≤x) : cumulative	P(X= x)
2	0.15	0.15
4	0.30	$0.30 - 0.15 = 0.15$
6	0.80	$0.80 - 0.30 = 0.50$
8	1	$1 - 0.80 = 0.20$
total	xxxxx	

x	P(X= x)
2	0.15
4	0.15
6	0.50
8	0.20
total	Total=1

Complement of probability:

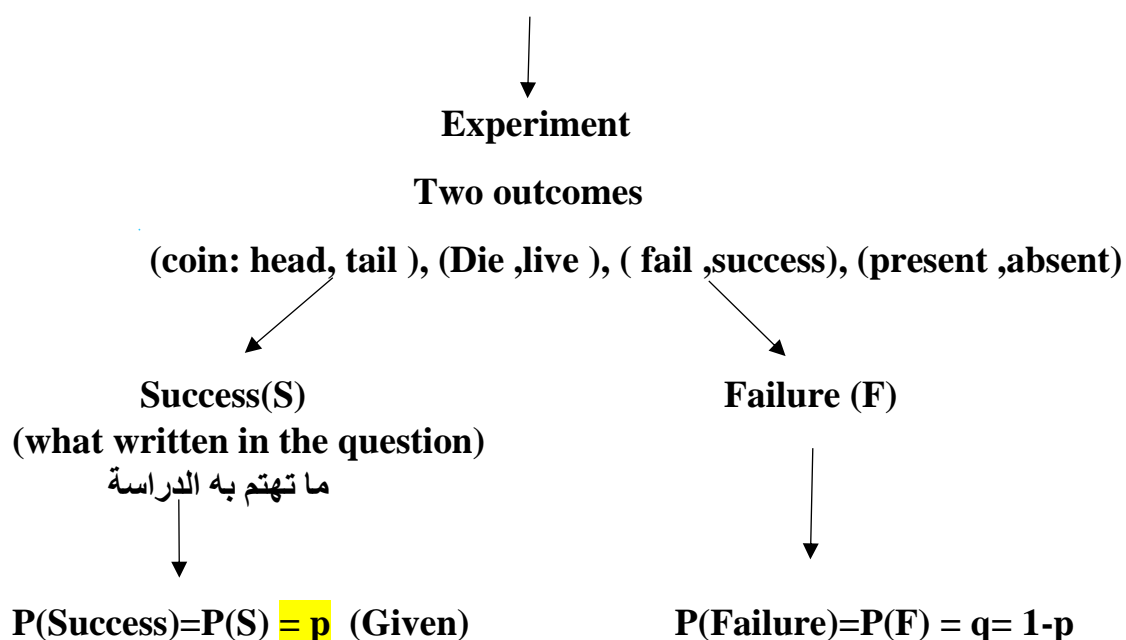
$$P(X > 1) = 1 - P(X \leq 1)$$

$$P(X \geq 1) = 1 - P(X < 1)$$

$$P(X \leq 1) = 1 - P(X > 1)$$

$$P(X < 1) = 1 - P(X \geq 1)$$

Binomial Distribution



* **n** : Sample size or number of repeating the experiment

(**n** given in the question)

$X \sim \text{Bin}(n, p)$

- **X** : Number of successes , ما تهتم به الدراسة

$x = 0, 1, 2, 3, 4, \dots, n$

- The parameter of Binomial distribution is n, p

- The mean is $\mu = n \times p$
- The variance is $\sigma^2 = n \times p \times q$
- The standard deviation is $\sigma = \sqrt{n \times p \times q}$

How to calculate probability in Binomial distribution

(By Calculator):

Example 1: Binomial distribution $n=5$, $p=0.3$, $q=0.7$

$$X \sim \text{Binomial}(5, 0.3)$$

Find

1) $P(X \leq 2) = ?$

$$\sum_{x=0}^2 ({}^5C_x * (0.3)^x * (0.7)^{(5-x)})$$

$\sum_{x=0}^2 ({}^5C_x * (0.3)^x * (0.7)^{(5-x)})$

To write the following in calculator :

Σ	\longrightarrow	Shift log
x	\longrightarrow	Alpha)
5C_2	\longrightarrow	5 shift \div 2

2) $P(X \geq 3) = ?$

$$\sum_{x=3}^5 ({}^5C_x * (0.3)^x * (0.7)^{(5-x)}) =$$

$\sum_{x=3}^5 ({}^5C_x * (0.3)^x * (0.7)^{(5-x)})$

Example : (Page 71)

Extra questions:

4. Find the probability that there will be more than or equal one snake bite cases in a month

$$\begin{aligned}P(X \geq 1) &= 1 - P(x < 1) \\&= 1 - P(X = 0) \\&= 1 - \frac{e^{-0.5}(0.5)^0}{0!} = 0.6065\end{aligned}$$

5. The mean of snake bite cases in a year

$$\mu = \lambda = 6$$

6. The variance of snake bite cases in a month

$$\sigma^2 = \lambda^* = \frac{\lambda}{12} = \frac{6}{12} = 0.5$$

7. The standard deviation of snake bite cases in 2 years

$$\sigma = \sqrt{\lambda^*} = \sqrt{2\lambda} = \sqrt{2(6)} = 3.4641$$

Sampling of \bar{X} (Page 86)

$$\text{Mean}(\bar{X}) = \mu$$

$$\text{Variance}(\bar{X}) = \frac{\sigma^2}{n}$$

$$\text{Standard deviation}(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

- Distribution of \bar{X} :

If population Normal or non-normal($n \geq 30$)

$$\bar{X} \sim \text{Normal}(\mu, \frac{\sigma^2}{n})$$

- Transform to Z or T (To use tables)

✚ If σ Known

* If σ unknown (use S instead), normal

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

$$T = \frac{\bar{X} - \mu}{s / \sqrt{n}}$$

How to use T-Table :

1) $t_{0.95} = ?$ (df=10)

2) $t_{0.90} = ?$ (df =12)

3) $t_{0.05} = ?$ (df=20)

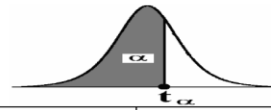
4) $t_{0.10} = ?$ (df=5)

طريقه اخرى لكتابة الاحتمال

1) $P(T < K) = 0.90$, (df =5)

2) $P(T < K) = 0.025$, (df =10)

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
90	1.2910	1.6620	1.9867	2.3685	2.6316
100	1.2901	1.6602	1.9840	2.3642	2.6259
120	1.2886	1.6577	1.9799	2.3578	2.6174
140	1.2876	1.6558	1.9771	2.3533	2.6114
160	1.2869	1.6544	1.9749	2.3499	2.6069
180	1.2863	1.6534	1.9732	2.3472	2.6034
200	1.2858	1.6525	1.9719	2.3451	2.6006
∞	1.282	1.645	1.960	2.326	2.576

Sampling of $(\bar{X}_1 - \bar{X}_2)$

Mean $(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2$, Variance $(\bar{X}_1 - \bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ Standard Deviation = $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

Distribution of \bar{X} :

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

•

Transform to Z or T (To use tables):

✚ If σ_1, σ_2 Known
Non-normal

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

* If σ_1, σ_2 unknown (but equal), Normal or
Normal (n_1, n_2 small)

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}$$

Where,

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 + n_2)}$$

Example (Page 99):

$$\begin{aligned} 1) \quad P(\hat{p} > 0.4) &= P\left(Z > \frac{0.4-0.45}{0.084}\right) = P(Z > -0.59) = 1 - P(Z < -0.59) \\ &= 1 - 0.2776 = 0.7224 \end{aligned}$$

$$\begin{aligned} 2) \quad P(0.4 < \hat{p} < 0.5) &= P\left(\frac{0.4-0.45}{0.084} < Z < \frac{0.5-0.45}{0.084}\right) = \\ &= P(-0.59 < Z < 0.59) = \\ &= 0.72240 - 0.27760 = 0.4448 \end{aligned}$$

Example : Page 102:

$$\begin{aligned} P(0.05 < \hat{P}_1 - \hat{P}_2 < 0.2) &= P\left(\frac{0.05-0.1}{0.06} < Z < \frac{0.2-0.1}{0.06}\right) = \\ &= P(-0.83 < Z < 1.67) = \\ &= 0.95254 - 0.20327 = 0.47827 \end{aligned}$$