

Tutorial 5

AID-521: Mathematics for Data Science

Abhishek Samantray, Vimal Kumar

September 2, 2021

Qn. 1.

A certain population has a normal distribution with parameters μ and σ^2 . The density is given by

$$f(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\}.$$

A random sample X_1, \dots, X_n is drawn from this population.

- (a) Find an estimator for μ using the method of moments.
- (b) Find an estimator for σ^2 using the method of moments.
- (c) Find the maximum likelihood estimator for μ , assuming μ is unknown and $\sigma^2 = \sigma_0^2$ is known.
- (d) Find the maximum likelihood estimator for σ^2 , assuming $\mu = \mu_0$ is known and σ^2 is unknown.

Ans/Sol.

Qn. 2.

For Mary's political poll, $n = 15$ voters were sampled. We wish to test

$$H_0 : p = 0.5$$

against the alternative hypothesis

$$H_a : p < 0.5.$$

The test statistic is Y , the number of sampled voters that favor Mary.

- (a) Calculate α if we select $RR = \{y \leq 2\}$ as the rejection region.
- (b) Suppose that she will receive 30% of the votes ($p = 0.3$). What is the probability β that the sample will erroneously lead us to conclude that H_0 is true and that she is going to win?

Ans/Sol.

Qn. 3.

A vice president in charge of sales for a large corporation claims that salespeople are averaging no more than 15 sales contacts per week. (He would like to increase this figure.)

As a check on his claim, $n = 36$ salespeople are selected at random, and the number of contacts made by each is recorded for a single randomly selected week. The mean and variance of the 36 measurements were 17 and 9, respectively.

- (a) Does the evidence contradict the vice president's claim? (Use a test with level $\alpha = 0.05$.)

Ans/Sol.

Qn. 4.

Historically, biology has been taught through lectures, and assessment of learning was accomplished by testing vocabulary and memorized facts. A teacher-developed new curriculum, Biology: A Community Content (BACC), is standards based, activity oriented, and inquiry centered. Students taught using the historical and new methods were tested in the traditional sense on biology concepts that featured biological knowledge and process skills. The results of a test on biology concepts were published in *The American Biology Teacher* and are given in the following table.¹

	Mean (μ)	Sample Size (n)	S.D. (σ)
Pretest: all BACC classes	13.38	372	5.59
Pretest: all traditional	14.06	368	5.45
Posttest: all BACC classes	18.50	365	8.03
Posttest: all traditional	16.50	298	6.96

- (a) Give a 90% confidence interval for the mean posttest score for all BACC students.
- (b) Find a 95% confidence interval for the difference in the mean posttest scores for BACC and traditionally taught students.
- (c) Does the confidence interval in part (b) provide evidence that there is a difference in the mean posttest scores for BACC and traditionally taught students? Explain.
- (d) Do the data provide support for the conjecture that the mean pretest scores do not differ for students subsequently taught using the two methods? Test using $\alpha = 0.01$.

Ans/Sol.

¹William Leonard, Barbara Speziale, and John Pernick, "Performance Assessment of a Standards-Based High School Biology Curriculum," *The American Biology Teacher* 63(5) (2001): 310–316.

Qn. 5.

We take a sample of size $n = 1$ from a uniform distribution defined on the interval $[0, \theta]$, where θ is unknown.

- (a) Find a 95% lower confidence bound for θ .

Ans/Sol.

Qn. 6.

The shopping times of $n = 64$ randomly selected customers at a local supermarket were recorded. The average and variance of the 64 shopping times were 33 minutes and 256 minutes², respectively.

- (a) Estimate μ , the true average shopping time per customer, with a confidence coefficient of $1 - \alpha = 0.90$.

Ans/Sol.

Qn. 7.

A psychological study was conducted to compare the reaction times of men and women to a stimulus. Independent random samples of 50 men and 50 women were employed in the experiment. The results are shown in the table below.

Men	Women
$n_1 = 50$	$n_2 = 50$
$\bar{y}_1 = 3.6$ seconds	$\bar{y}_2 = 3.8$ seconds
$s_1^2 = 0.18$	$s_2^2 = 0.14$

- (a) Do the data present sufficient evidence to suggest a difference between true mean reaction times for men and women? Use $\alpha = 0.05$.

Ans/Sol.

Qn. 8.

Four groups of students were subjected to different teaching techniques and tested at the end of a specified period of time. As a result of dropouts from the experimental groups (due to sickness, transfer, etc.), the number of students varied from group to group. The data is shown in the table below.

	1	2	3	4
	65	75	59	94
	87	69	78	89
	73	83	67	80
	79	81	62	88
	81	72	83	
	69	79	76	
		90		
T_i	454	549	425	351
n_i	6	7	6	4
\bar{T}_i	75.67	78.43	70.83	87.75

- (a) Do the data present sufficient evidence to indicate a difference in mean achievement for the four teaching techniques?

Ans/Sol.