Mid-term Exam

DATA ANALYSIS A - FOY28 A a.a. 2025-26

February 9, 2026

Ex. 1 Distribution (10 min)

Given the following set of measurements of a physical quantity x:

x	5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9
Frequency	1	3	5	7	9	7	5	3	1

a) Draw the distribution

b) Calculate the mean and the median

c) Calculate the variance σ^2 and the standard deviation σ.

d) Calculate the error on the mean σ_μ

SOLUTION

b) $\mu = 5.5$; $median = 5.5$; c) $\sigma^2 = 0.034146$; $\sigma = 0.18478$; d) $\sigma_\mu = 0.02885$.

Ex. 2 Weighted mean (5 min)

From the following 5 measurements of the speed of light c, calculate the mean and its uncertainty:

$$299794000 \pm 3000\ m/s$$

$$299791000 \pm 5000\ m/s$$

$$299770000 \pm 2000\ m/s$$

$$299789000 \pm 3000\ m/s$$

$$299790000 \pm 3000\ m/s$$

SOLUTION

Weighted mean. $\mu_c = 299782577\ m/s$; $\sigma_c = 1266\ m/s \rightarrow c = (299782577 \pm 1266)\ m/s$

Ex. 3 Significant figures (4 min)

Report the following results correctly, by stating the uncertainty to only 1 significant digit.

a) $699.06133000 \pm 0.056291$

b) 20511 ± 0.5102

c) 30.9 ± 11

d) 31.25 ± 0.034953

e) 1261.29 ± 200

f) 0.0000139 ± 0.0000024

SOLUTION

a) 699.06 ± 0.06; b) 20511.0 ± 0.5; c) $31 \pm 11 \approx (3 \pm 1) \cdot 10$; d) 31.25 ± 0.03; e) $1261 \pm 200 \approx (1.3 \pm 0.2) \cdot 10^3$; f) $0.000014 \pm 0.000002 = (1.4 \pm 0.2) \cdot 10^{-5}$

Ex. 4 Covariance and Correlation (10 min)

Consider the following data. Calculate the correlation coefficient between variables x and y. What can you say about x and y ?

x	0	2	4	5	7	8	10
y	4	3	1	1	2	1	0

SOLUTION

$\mu_X = 5.1428$; $\mu_Y = 1.71428$; $\sigma_X = 3.2261$; $\sigma_Y = 1.2777$; $cov(X, Y) = -3.5306$; $\rho = -0.8564$

Ex. 5 (Poisson, 3 min)

The number of typing mistakes made by a secretary has a Poisson distribution. The mistakes are made independently at an average rate of 1.65 per page. Find the probability that a three-page letter contains no mistakes.

SOLUTION

If $\lambda = 1.65\ err/page$, for three pages the average number of errors are $\lambda = 3 \cdot 1.65 = 4.95$. We use the Poisson distribution, since we are not in the case of a success/failure experiment and the probability of committing a mistake is not defined (and likely not constant). $p(k = 0) = \frac{\lambda^k \cdot e^{-\lambda}}{k!} = \frac{\lambda^0 \cdot e^{-\lambda}}{0!} = e^{-4.95} = 0.007083$.

Ex. 6 Uniform distribution (4 min)

The amount of time, in minutes, that a person must wait for a bus is uniformly distributed between zero and 15 minutes, inclusive.

a) Draw the probability distribution

b) What is the average waiting time and what is the standard deviation ?

c) What is the probability that a person waits less than 12.5 minutes?

d) What is the probability that a person waits more than 6 minutes and less than 9 minutes ?

SOLUTION

b) $\mu = \frac{b-a}{2} = 15/2 = 7.5\ min$; $\sigma = \frac{b-a}{\sqrt{12}} = 15/\sqrt{12} = 4.33$. c) $P(x < 12.5) = 1/15 \cdot 12.5 = 0.833$. d) $P(6 < x < 9) = 1/15 \cdot 3 = 0.2$.

Ex. 7 Poisson (10 min)

A 5-litre bucket of water is taken from a swamp. The water contains 75 mosquito larvae. A 200mL flask of water is taken form the bucket for further analysis.

a) What is the expected number of larvae in the flask?

b) What is the probability that the flask contains at least one mosquito lava?

c) What is the maximum number of larvae I can expect to find (i.e. the maximum or less) in the flask with a total probability not exceeding 70%?

SOLUTION

a) If in 5 liters, the expected number of larvae is $\lambda = 75$; in 200mL $\lambda = 75/5 \cdot 0.2 = 3$.
b) $P(k \geq 1) = 1 - P(k = 0, \lambda = 3) = 1 - \frac{\lambda^0 \cdot e^{-\lambda}}{0!} = 1 - e^{-3} = 0.9502$.
c) $P(0) = 0.04978$, $P(1) = 0.14936$; $P(2) = 0.22404$; $P(3) = 0.22404$; $P(4) = 0.16803$. $P(0) + P(1) + P(2) + P(3) < 0.7$, while $P(0) + P(1) + P(2) + P(3) + P(4) > 0.7$. I can conclude that there is a probability of less than 70% of finding $k \leq 3$ larvae (maximum number of larvae expected with a total probability not exceeding 70% is 3).

Ex. 8 (Error propagation, 6-7 min)

Ohm's law states that the electric current I passing through a resistance R is directly proportional to the voltage V across the resistance: $V = R \cdot I$. Calculate the resistance R and its uncertainty if you measure a value of current $I = (5.0 \pm 0.1)A$ and you provide a voltage $V = (284 \pm 7)V$. Approximate the result correctly (1 significant figure on the error)
The units of resistance are ohms, $\Omega = \frac{volts(V)}{ampere(A)}$.

SOLUTION

$R = \frac{V}{I} = \frac{284}{5} = 56.8\Omega$
$\frac{\sigma_R}{R} = \sqrt{\left(\frac{\sigma_V}{V}\right)^2 + \left(\frac{\sigma_I}{I}\right)^2} = \sqrt{\left(\frac{7}{284}\right)^2 + \left(\frac{0.1}{5}\right)^2} = \sqrt{0.000607518 + 0.0004} = 0.03174143$, $\sigma_R = 1.8029\Omega$
The resistance value is $R = (57 \pm 2)\Omega$

Ex. 9 (Binomial, 5-8min)

The probability that a driver must stop and pay a bribe to corrupt police officers at a traffic light along Corruption Road is 0.2. There are 15 of traffic lights on the unfortunate road.

a) What is the probability to pass unharmed ?

b) What is the probability that you have to pay 2 bribes ?

c) What is the probability that you will be stopped at 1 or more times ?

SOLUTION

a) $P(k = 0, p = 0.2, n = 15) = 0.2^0 \cdot 0.8^{15} \cdot \frac{15!}{0!15!} = 0.8^{15} \approx 0.035 \approx 3.5\%$

b) $P(k = 2, p = 0.2, n = 15) = 0.2^2 \cdot 0.8^{13} \cdot \frac{15!}{2!13!} \approx 0.2309 \approx 23\%$

c) $P(k \geq 1) = 1 - P(0) \approx 1 - 0.035 = 0.965 \approx 96.5\%$

Ex. 10 (Gaussian, 15min)

The finish times for marathon runners during a race are normally distributed with a mean of 195 minutes and a standard deviation of 25 minutes.

a) What is the probability that a runner will complete the marathon within 3 hours?

b) Calculate to the nearest minute, the time by which the first 8% runners have completed the marathon.

c) What proportion of the runners will complete the marathon between 3 hours and 4 hours?

SOLUTION

a) $P(x < 180)$? $z = \frac{x-\mu}{\sigma} = \frac{180-195}{25} = -0.6$. $P(-0.6 < z < 0.6) = 0.4515$.
$P(x < 180) = \frac{1-0.4515}{2} = 0.27424$.

b) What is $\bar{x}$ so that $P(x < \bar{x}) = 0.08$. $P(-\bar{z} < z < \bar{z}) = 1 - 0.08 \cdot 2 = 0.84 \rightarrow \bar{z} = -1.4$ (negative since it lays on the left of the mean). $\bar{z} = \frac{\bar{x}-\mu}{\sigma}$, $\bar{x} = \bar{z} \cdot \sigma + \mu = -1.4 \cdot 25 + 195 = 160$ min.

c) $x_1 = 180$, $x_2 = 240$. $z_1 = \frac{x_1-\mu}{\sigma} = \frac{180-195}{25} = -0.6$; $z_2 = \frac{x_2-\mu}{\sigma} = \frac{240-195}{25} = 1.8$.
$P(z_1 < z < 0) = \frac{0.4515}{2} = 0.22575$; $P(0 < z < z_2) = \frac{0.9281}{2} = 0.46405$.
$P(180 < x < 240) = 0.22575 + 0.46405 = 0.6898$.