

Bank Soal**Statistics 2**

1. Given that $X \sim \text{Po}(2.1)$, find
 - (a) $\mathbb{P}(U = 2)$
 - (b) $\mathbb{P}(U < 3)$
 - (c) $\mathbb{P}(U \geq 4)$
2. The number of deaths in a village has a mean 3 per week. Assuming that the number of deaths follows a Poisson distribution, calculate the probability that
 - (a) there are less than 2 deaths in a particular week.
 - (b) there are no deaths in a particular week.
3. The probability that a patient has a particular disease is 0.005. One day 100 people go to their doctor.
 - (a) What is the probability that exactly 2 of them have the disease?
 - (b) What is the probability that 3 or more of them have the disease?
4. Sales of cooking oil bought in a shop during a week follow a Poisson distribution with mean 80. How many units should be kept in stock to be at least 80% certain that supply will be able to meet demand?
5. The variables W, X, Y and Z have means 1,2,3 and 4 respectively and variances 4,3,2 and 1 respectively. Find
 - (a) $\mathbb{E}(W + X + Y - 1.5Z)$ and $\text{Var}(W + X + Y - 1.5Z)$
 - (b) $\mathbb{E}(2W + 4X + 3Y - Z)$ and $\text{Var}(2W + 4X + 2Y - Z)$
6. The heights of a population of male students are distributed normally with mean 170 cm and standard deviation 25 cm. The heights of a population of female students are distributed normally with mean 161 cm and standard deviation 9 cm. Find the probability that a randomly chosen female is taller than a randomly chosen male.
7. A factory distributes boxes of TV and boxes of computer, each boxes containing exactly one object. It is known that the mass of the computers are normally distributed with mean 2.1 kg and standard deviation 0.4 kg and the mass of the TVs are normally distributed with mean 1.9 kg and variance 0.7 kg². Find the probability the mass of five computer boxes differs from the mass of three TV boxes by at most 1 kg.
8. A random variable X has mean 1 and variance $\frac{3}{4}$. Let $Y = ax + b$. If $\mathbb{E}(Y) = 10$ and $\text{Var}(Y)=12$, find the value of $a - b$.
9. Given the probability density function

$$f(x) = \begin{cases} k(27 - x^3) & \text{for } |x| < 3 \\ 0 & \text{otherwise.} \end{cases}$$

- (a) Find the constant k .
- (b) Find the median value of X .
- (c) Find $\mathbb{P}(X \geq 1)$.
- (d) Explain why the following function isn't suitable as a probability density function.

$$f(x) = \begin{cases} \frac{1}{216}(27 - x^3) & \text{for } |x| < 4 \\ 0 & \text{otherwise.} \end{cases}$$

10. The random variable X has probability density function

$$f(x) = \begin{cases} ce^{-cx} & \text{for } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

- (a) Verify that this is a properly defined probability density function, regardless of the value of the constant c .
- (b) Find the median value of X in terms of c .
- (c) Suppose $c = 0.7$ and the above distribution is proposed as a model for the working length, in years, of a Wardaya College staff. Find the expected number out of 100 staffs of Wardaya College that would work in a class interval of 1-2 years.
- (d) Find $\text{Var}(X)$ given $c = 0.6$.

11. The continuous random variable X has probability density function

$$f(x) = \begin{cases} (x-a)(2x-5a)(x-3a) & \text{for } a \leq x \leq 2a \\ 0 & \text{otherwise.} \end{cases}$$

- (a) Verify that the following function is a properly defined probability density function.
- (b) Show that $a^4 = \frac{6}{7}$.

12. The continuous random variable X has probability density function

$$f(x) = \begin{cases} \frac{k}{x} & \text{for } 1 \leq x \leq a \\ 0 & \text{otherwise.} \end{cases}$$

- (a) Show that $k = \frac{1}{\ln a}$.
- (b) Find the median value of X in terms of a .
- (c) Find $\mathbb{E}(X)$ and $\text{Var}(X)$ given $a = 4$.
- (d) Suppose that $a = 3$ and X is proposed to be a model for the length life of a species of lion. Find the probability that a lion lasts between 1.25 and 2.5 years only.
- (e) Let Y be an event where a lion lasts between 1.25 and 2.5 years only. Find $\frac{\mathbb{P}(Y|a=3)}{\mathbb{P}(Y|a=2)}$.

13. Give a reason why the following procedure will not give a random sample of the letters of the alphabet.
Repeatedly choose a page at random from a phone book and take the initial letter of the first contact person defined on the page. Describe a suitable way of obtaining such a sample.
14. A rectangular field is gridded into squares of side 1 cm. At one time of the year the number of worms in the field can be modelled by a Poisson distribution with mean 3 per cm^2 .
- A random sample of 140 squares is observed and the number of worms in each square is counted. Find the probability that the sample mean number of worms is at most 1.8.
 - The concentration of phosphorus in the field can be modelled by a normal distribution with mean 15% and standard deviation 3%. A random sample of 100 squares is observed and the concentration of phosphorus is measured. Find the probability that the sample mean concentration of phosphorus is at least 13%.
15. The mean of a random sample of n observations drawn from a $N(\mu, \sigma^2)$ distribution is denoted by \bar{X} . Given that $\mathbb{P}(|\bar{X} - \mu| > 0.4\sigma) < 0.1$.
- Find the smallest value of n
 - With the same value of n , find $\mathbb{P}(\bar{X} < \mu + 0.2\sigma)$
 - With the same value of n , find $\mathbb{P}(\mu - 0.1\sigma \leq \bar{X} \leq \mu + 0.25\sigma)$
16. A random sample of $3n$ observations is taken of the random variable $X \sim B(n, p)$ and $p > 0.5$. The sample mean is denoted by \bar{X} . It is given that $\text{Var}(\bar{X}) = 0.07$ and $E(\bar{X}) = 70$.
- Find the values of p and n .
 - Find $\mathbb{P}(\bar{X} > 71)$.
 - Another random sample of n observations is taken from X . Find the exact value of the new expectation and variance of the sample.
17. The following table is the sample observations of the mass of 60 grade 12 students.

Mass (kg)	f
54	6
55	10
56	14
57	20
58	10

- With 97% confidence level, find the interval of the population mean mass of the grade 12 students.
- Explain why your previously obtained result is just an estimate.

18. The contents of 200 jars of peanut butter selected randomly from a package delivered to a market are weighed and the results, x grams, summarised by $\Sigma(w - 300) = -21$ and $\Sigma(w - 400)^2 = 1024$.
- Calculate unbiased estimates of the package mean and variance of the mass of the peanut butter in a jar.
 - Calculate a symmetric 95% confidence interval for the jar mean mass.
19. A biased 4-sided dice was thrown 300 times and 89 fours were obtained. Calculate a symmetric 97% confidence interval of p , the probability of obtaining a four in a single throw of the dice.
Hence, estimate the smallest extra number of times the dice should be thrown for the width of the symmetric 97% confidence interval of p to be at most 0.075.
20. A random sample of 12 people studying in Wardaya College with 300 students are asked, at a certain day, how tall they are in metres. The results in metres are as follow.

1.56, 1.67, 1.61, 1.81, 1.72, 1.63, 1.65, 1.54, 1.49, 1.73, 1.81, 1.70

Calculate unbiased estimates of the mean and variance of the height of all students studying there.

21. A factory received complaints from consumers for the net weight of a canned food X isn't the same as what is written on the can seal, which is 10 ounces. A research is then conducted by taking 21 random samples of canned food X . It is known that the cans have mean weight 9.8 ounces with standard deviation 0.4 ounces. With significance level of 0.05, test where the complaints are justified or not. State the null and the alternative hypotheses.
22. A cutting machine is used to cut metal with a specific size. The factory states that the metal produced has length 50 cm. To find the truth, a research is conducted on 100 random samples of metal plate. It is then known that the mean length of the plates are 50.3 cm with standard deviation 0.5 cm. Assuming that the lengths of the plates follow a normal distribution, find out whether the size of the produced metal has changed or not, given a 0.05 significance level. State the null and the alternative hypotheses as well.
23. A student in Wardaya College states that the height of Wardaya College students is normally distributed with mean 165 cm and standard deviation 2 cm. A random sample of 12 students are asked about their height and the results are as follows.

1.56, 1.67, 1.61, 1.81, 1.72, 1.63, 1.65, 1.54, 1.49, 1.73, 1.81, 1.70

Test, at the 5% significance level, whether the student's statement is justified.

24. The contents of 200 jars of peanut butter selected randomly from a package delivered to a market are weighed and the results, x grams, summarised by $\Sigma(w - 300) = -210$ and $\Sigma(w - 400)^2 = 1024$. It is known that in regular the mean mass of the jars in a single package is 298.9 grams. Test, at the 2.5% significance level, whether these

results indicate that the reported mean is too high.

Suppose the mean mass of the jars is 299.0 grams. With the same sample, find the p-value of a test of whether the reported mean of 299.0 grams is too high.

25. (Miscellaneous exercise 6, page 120)

An ambulance station serves an area which includes more than 10 000 houses. It has been decided that if the mean distance of the houses from the ambulance station is greater than 10 miles then a new ambulance station will be necessary. The distance, x miles, from the station of each of a random sample of 200 houses was measured, the results being summarised by $\Sigma x = 2092.0$ and $\Sigma x^2 = 24994, 5$.

(a) Calculate, to 4 significant figures, unbiased estimates of

- i. the population mean distance, μ miles, of houses from the station.
- ii. the population variance of the distance of the houses from the station.

State what you understand by 'unbiased estimate'

(b) A test of the null hypothesis $\mu = 10$ against the alternative hypothesis $\mu > 10$ is carried out at the $\alpha\%$ significance level, using a random sample of size 200. The rejection region for this test is $\bar{X} \geq 10.65$, where \bar{X} denotes the sample mean.

- i. Calculate the value of α .
- ii. State the conclusion of test using the sample data.

(c) Suppose that it could not be assumed that the distances are normally distributed. State whether the answers to part (a) and part (b) would still hold.

26. (Exercise 7B, page 127)

The drop-out rate of students enrolled at a certain university is reported to be 13.2%. The Dean of Students suspects that the drop-out rate for science students is greater than 13.2%, and she examines the records of a random sample of 95 of these students. The number of drop-outs was found to be 20. Test the Dean's suspicion at the 2% significance level.

27. The number of rotten avocados in a box harvested from grandma's garden has a Poisson distribution with mean λ . One day I received a box with 4 rotten avocados in it. With a 10% significance level, calculate the least value of λ such that I can conclude that the avocados are harvested from grandma's garden, correct to 4 decimal places.

END OF PAPER