Sample Final 3

Instructions:

1. Questions 1 to 26 are short answer questions. On the real final exam, put your answer in the box provided. An example is given in question 1. **You must show some work for each question in order to receive any marks**.
2. Questions 27, 28 and 29 are full-answer questions. For full-answer questions, marks will be deducted for incomplete or poorly presented solutions.
3. Space will be provided for you to work out your answers on the actual exam.

Duration: You should be able to complete this midterm within 3 hours. If you cannot, this means that more practice is still needed.

Questions 1, 2 and 3 refer to the following scenario: The 20 houses in a certain city block were classified as in the following table. Suppose one of these 20 houses is selected at random.

	Number of Stories			
	1	2	3	
Owner Occupied	2	1	5	
Rented	5	4	3	

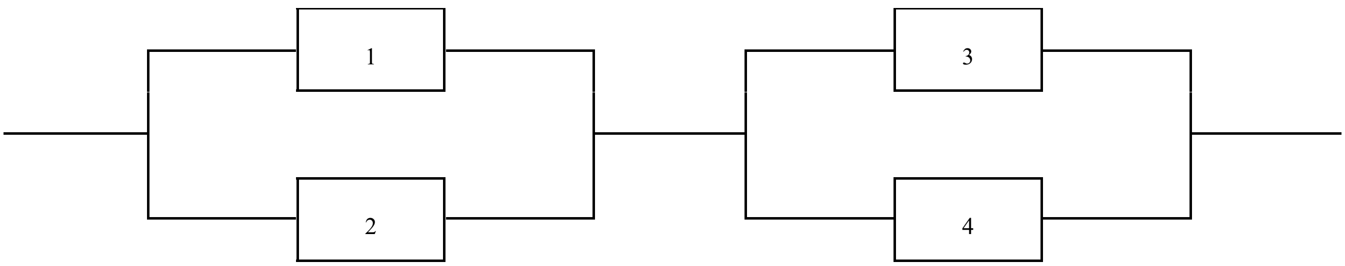
1. What is the probability that the selected house is neither rented nor has fewer than two stories?

Answer:

2. If the selected house is rented, what is the probability that it has fewer than three stories?

3. Which of the following three statements are true?
(i) The events 'selected house is rented' and 'selected house has an odd number of stories' are mutually exclusive.
(ii) The events 'selected house is rented' and 'selected house has an odd number of stories' are independent.
(iii) The events 'selected house has an odd number of stories' and 'selected house has an even number of stories' are independent.

Questions 4, 5 and 6 refer to the following scenario: Consider the system of components connected as in the following diagram. The subsystem consisting of components 1 and 2 works if either 1 or 2 works. The subsystem consisting of components 3 and 4 works if either 3 or 4 works. The whole system works if both subsystems work. The probabilities that components 1 through 4 work are .1, .2, .3, .4, respectively. Assume the four components function independently.



4. What is the probability that the subsystem consisting of components 3 and 4 works?

5. What is the probability that the whole system works?

6. What is the expected number of components that work in the subsystem consisting of components 3 and 4?

7. A box contains 5 coins, of which two are honest and three are two-headed. One coin, selected at random from the box, is tossed and observed to come up heads. What is the probability that the selected coin is two-headed?

8. Let A and B be two events such that $P(A|B) = 0.33$, $P(B|A) = 0.55$, and $P(A \cup B) = 0.66$. Find $P(A \cap B)$.

9. In a certain factory, accidents occur at random times at the average rate of 4 accidents every 6 working days. What is the probability that at least 2 accidents occur in the next 3 working days?

10. Suppose that 40% of adult males are overweight. In a random sample of 20 adult males what is the probability that between 8 and 10 (inclusive) are overweight?

Questions 11, 12, 13 refer to the following scenario: Let X be a discrete random variable with the following probability distribution:

x	-2	-1	0	1	2
$P(X = x)$	0.04	0.12	0.36	0.28	0.20

11. Find the expected value of X.
12. Find the variance of X.
13. Find $E(2X^3 + 6)$.

Questions 14, 15 and 16 refer to the following scenario: The yield X (in litres) of a certain process is a continuous random variable with probability distribution given by the following probability density function:

$$f(x) = \begin{cases} \frac{x^3}{20} & \text{if } 1 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

14. Find the 10-th percentile of this distribution.
15. Find the standard deviation of X.
16. A random sample of 10 observations is to be drawn from the distribution of X. What is the expected number of these observations that will fall below 2 litres?

Questions 17, 18 and 19 refer to the following scenario: Two random variables X and Y have the following joint probability mass function:

			y	
	$p(x, y)$	-2	0	2
	-1	.08	0	.02
x	0	0	.20	.40
	1	.12	0	.18

17. Find the variance of Y.
18. Find the covariance of X and Y.
19. Find $P(X \geq Y | X + Y > -2)$.
20. Suppose the lifetime, X, of a certain type of component is a random variable having an exponential distribution with mean lifetime $\mu = \frac{1}{\lambda} = 5$ years. What is the probability that a two-year-old component of this type will last an additional 4 years?

Questions 21 and 22 refer to the following scenario: The amount of fill, X, for a certain size box of Brand K cereal is a normally distributed random variable with mean $\mu = 750$ grams and standard deviation $\sigma = 4$ grams.

21. What proportion of such cereal boxes will contain more than 755 grams of fill?

22. Find w such that only 20% of these cereal boxes contain less than w grams of fill.

23. Suppose X_1, X_2, X_3, X_4 is a random sample of size $n = 4$ from a population distribution having unknown mean μ and unknown standard deviation σ. Consider the following three estimators for μ.

$$R = \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3; \qquad U = \frac{1}{10}X_1 + \frac{2}{10}X_2 + \frac{3}{10}X_3 + \frac{4}{10}X_4; \qquad W = \frac{1}{5}X_1 + \frac{1}{5}X_2 + \frac{1}{5}X_3 + \frac{2}{5}X_4.$$

Which of the following statements are true?

i The variance of W is less than the variance of U, i.e. $Var(W) < Var(U)$.

ii The variance of W is less than the variance of R, i.e. $Var(W) < Var(R)$.

iii The variance of U is less than the variance of R, i.e. $Var(U) < Var(R)$.

Questions 24 and 25 refer to the following scenario: A survey was conducted to estimate the proportion p of passengers on a certain airline who are dissatisfied with the food served. A random sample of 200 passengers using this airline was drawn, and, of these, 42 were dissatisfied with the food.

24. Compute the upper limit of a 90% confidence interval for p.

25. Use these data as a pilot study to determine the number of additional observations needed to estimate p within 4 percentage points with 95% confidence (i.e. 95% CI width = .08).

26. A random sample of 50 measurements of arsenic in copper yielded the following data (in percent):
Sample Mean = .180 percent; Sample Standard Deviation = .026 percent
Compute the lower limit of a 97% confidence interval for the true mean percent of arsenic in the sampled copper.

27. LONG ANSWER QUESTION: An experiment was performed to compare two overnight parcel delivery services. Each delivery service was given identical parcels at the same time for the same destination on each of 6 occasions. The delivery times in hours are in the following table. Let μ_1 and μ_2 denote the true mean delivery times for Delivery Service I and Delivery Service II, respectively; and let $\mu_D = \mu_1 - \mu_2$. Assume the relevant population distribution(s) is(are) normal.

Occasion	1	2	3	4	5	6
Service I	13	15	15	11	11	15
Service II	10	10	15	8	12	11

(a) Define the population parameter(s) of interest.

(b) State the null and alternative hypotheses in terms of the parameter.

(c) State the test statistic you will use. What distribution (including degrees of freedom, if appropriate) will you use to calculate the p-value.

(d) Find the observed value of the test statistic.

(e) Compute (or bracket) the p-value within the accuracy of the tables.

(f) What level of evidence against H_0 do you find?

28. LONG ANSWER QUESTION: At a certain gasoline station, the amount of gasoline sold to a randomly selected customer has mean 20 litres and standard deviation 5 litres.

    (a) On a day when 300 customers independently purchase gasoline at this station, what is the probability that the total amount of gasoline sold exceeds 6100 litres?

    (b) What is the probability that on a day with 300 customers, the total amount of gasoline sold is less than the total amount sold on a day with 290 customers?

29. LONG ANSWER QUESTION: A certain brand of microwave oven was priced at random samples of stores in Toronto and in Vancouver with the tabulated results. Let μ_1 denote the true mean price of these microwave ovens in Toronto, and let μ_2 denote the true mean price in Vancouver. Assume the relevant population distribution(s) is(are) normal. Do these data provide evidence that the true mean price is higher in Vancouver than in Toronto?

Toronto	$m = 5$	$\bar{x} = 470$	$s_1 = 80$
Vancouver	$n = 7$	$\bar{y} = 560$	$s_2 = 50$

(a) Define the population parameter(s) of interest.

(b) State the null and alternative hypotheses in terms of the parameter.

(c) State the test statistic you will use. What distribution (including degrees of freedom, if appropriate) will you use to calculate the p-value.

(d) Find the observed value of the test statistic.

(e) Compute (or bracket) the p-value within the accuracy of the tables.

(f) What level of evidence against H_0 do you find?

(g) Construct a 90% confidence interval for $\mu_1 - \mu_2$.

Answers :

1. 0.3	2. 0.75	3. None
4. 0.58	5. 0.1624	6. 0.70
7. 0.75	8. 0.1715	9. 0.594
10. 0.456	11. 0.48	12. 1.1296
13. 8.88	14. 1.732	15. 0.4585
16. 1.875	17. 2.56	18. 0.08
19. 0.3478	20. 0.449	21. 0.1056
22. 746.64	23. all	24. 0.2574
25. 199	26. 0.172	

27.(a) μ_D = true mean difference, Service I minus Service II
(b) $H_0 : \mu_D = 0$ and $H_1 : \mu_D \neq 0$
(c) $T = \frac{\bar{D}-0}{s_D/\sqrt{n_d}} \sim t_{(5)}$
(d) $t_{obs} = 2.44$
(e) $p-value = 2 * P(T_{(4)} \geq 2.44)$, $0.05 < p-value < 0.10$
(f) There is moderate evidence against the null hypothesis.

28.(a) Let X_i = the amount of gasoline sold to the $i-th$ customer on a 300-customer day
$T_0 = X_1 + X_2 + \cdots + X_{300}$, $E(T_0) = 300 * 20 = 6000$, $Var(T_0) = 300 * 25 = 7500$
$T_0 \sim N(6000, 7500)$ by the CLT
$P(T_0 > 6100) \approx P\left(Z > \frac{6100-6000}{\sqrt{7500}}\right) = P(Z > 1.15) = 0.1251$
(b) Let Y_i = the amount of gasoline sold to the $i-th$ customer on a 290-customer day
$W_0 = Y_1 + Y_2 + \cdots + Y_{290}$, $E(W_0) = 290 * 20 = 5800$, $Var(W_0) = 290 * 25 = 7250$
$T_0 - W_0 \approx N(6000 - 5800 = 200, 7500 + 7250 = 121.45^2)$
$P(T_0 - W_0 < 0) \approx P\left(Z < \frac{0-200}{121.45}\right) = P(Z < -1.65) = 0.0495$

29.(a) μ_1 = true mean price in Toronto, $\mu 21$ = true mean price in Vancouver,
We are interested in $\mu_1 - \mu_2$
(b) $H_0 : \mu_1 - \mu_2 = 0$, $H_1 : \mu_1 - \mu_2 < 0$.
(c) Test statistic, since $\frac{s_1}{s_2} = 1.6 > 1.4$, use unpooled t-test

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{s_1^2/m + s_2^2/n}} \sim \text{Student}(\nu)$$

$$\nu = \text{integer part} \left[\frac{(s_1^2/m + s_2^2/n)^2}{\frac{(s_1^2/m)^2}{m-1} + \frac{(s_2^2/n)^2}{n-1}} \right]$$

(d) Observed test statistic -2.22 and $\nu = 6$.
(e) $p-value = P(T_{(6)} \leq -2.22) = P(T_{(6)} \geq 2.22)$; $0.025 < p-value < 0.5$
(f) There is **strong** evidence against H_0.
(g) 90% CI is: $470 - 560 \pm (1.943)(40.46)$ or $(-168.61, -11.39)$