DISCLAIMER: This is a sample and your final exam might be easier or harder than this exam. It is a good idea to go through your sample midterms, actual midterms, assignments and problems from the course notes. This sample final is short to give you an idea of the format of the exam. The actual exam is 2.5 hrs long and will consist of more questions.

Section A: For multiple-choice questions 1-11, **circle** the best answer. There is only one correct answer to each question.

1. Random variable X is normally distributed, with a mean of 25 and a standard deviation of 4. Which of the following is the appropriate interquartile range for this distribution?

   A. 27.20 - 22.80 = 4.40

   B. 27.70 - 22.30 = 5.40

   C. 28.70 - 21.30 = 7.40

   D. 2.00(4.00) = 8.00

   E. 1.50(4.00) = 6.00

2. A bag contains exactly three balls numbered 2, 4 and 6 respectively. Random samples of two balls are drawn from the bag with replacement. The average, $\bar{X} = \frac{X_1+X_2}{2}$, where X_1 and X_2 are the numbers on the selected balls, is recorded after each drawing. Which of the following describes the sampling distribution of $\bar{X}$?

   A.

$\bar{X}$	2	3	4	5	6
Probability	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

   B.

$\bar{X}$	2	3	4	5	6
Probability	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{3}$	$\frac{2}{9}$	$\frac{1}{9}$

   C.

$\bar{X}$	2	3	4	5	6
Probability	0	0	1	0	0

   D.

$\bar{X}$	2	3	4	5	6
Probability	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{1}{10}$

   E. Normal distribution with mean 4 and standard deviation $\sqrt{2}$

3. Events A and B are independent. If $P(A) = \frac{1}{3}$, which of the following statements is true?

   A. If $P(B) = \frac{4}{5}$, then $P(A \cap B) = \frac{7}{15}$.
   B. If $P(B) = \frac{1}{2}$, then $P(B \mid A) = \frac{1}{6}$.
   C. If $P(B) = \frac{2}{3}$, then $P(A \cup B) = \frac{2}{9}$.
   D. If $P(B) = \frac{1}{4}$, then $P(A \cap B) = \frac{5}{12}$.
   E. If $P(B) = \frac{1}{5}$, then $P(A \mid B) = \frac{1}{3}$.

4. The table below represents the decision based on a sample. Which of the following best describes the entries designated by (2) and (3) respectively?

		Truth about the population	
		H_0 true	H_a true
Decision based on sample	Reject H_0	(1)	(2)
	Do not reject H_0	(3)	(4)

   A. Type I, Type II
   B. Correct decision, Correct Decision
   C. Type II, Type I
   D. Correct decision, Type I
   E. Type I, Correct decision

5. An exam consists of 40 multiple-choice questions to each of which five answers are given, only one of which is correct. For each correct answer, a candidate gets 1 mark, and no penalty is applied for getting an incorrect answer. A particular candidate answers each question purely by guess-work. What is the estimated probability this student obtains a score greater than or equal to 10?

   A. 0
   B. 0.2776
   C. 0.4207
   D. 0.5793
   E. 0.7224

6. The number of kWh of electricity consumed on a randomly chosen day in a small factory can be modelled by the random variable X with probability density function given by

$$f(x) = \begin{cases} \frac{75}{x^2} & 50 \leq x \leq 150 \\ 0 & \text{otherwise.} \end{cases}$$

Find the probability that on a randomly chosen day more than 100 kWh of electricity will be consumed.

A. 0.10

B. 0.25

C. 0.50

D. 0.75

E. 0.90

7. X is uniformly distributed over the interval [-1, 1]. Find the probability density function of $Y = e^X$.

A.

$$f_Y(y) = \begin{cases} 0.0729e^y & \frac{1}{e} \le y \le e \\ 0 & \text{otherwise.} \end{cases}$$

B.

$$f_Y(y) = \begin{cases} \frac{1}{2y} & \frac{1}{e} \le y \le e \\ 0 & \text{otherwise.} \end{cases}$$

C.

$$f_Y(y) = \begin{cases} \frac{2e^2}{e^4-1}y & \frac{1}{e} \le y \le e \\ 0 & \text{otherwise.} \end{cases}$$

D.

$$f_Y(y) = \begin{cases} \frac{e}{1-e^2} & \frac{1}{e} \le y \le e \\ 0 & \text{otherwise.} \end{cases}$$

E.

$$f_Y(y) = \begin{cases} \frac{1}{2}ln y + \frac{1}{2} & \frac{1}{e} \le y \le e \\ 0 & \text{otherwise.} \end{cases}$$

8. A random variable X has the probability distribution given in the following table.

x	2	3	4	5
P(X = x)	0.1	0.2	0.3	0.4

Find $E(|x-4|)$

A. 0

B. 0.6

C. 0.8

D. 1

E. 4

9. It can get very hot for the lecturer on stage. Not only do they have to conduct a lecture, but lecture theatres spotlights are usually bright and intense. Independent random samples of spotlights were obtained from four different lecture halls. The wattage of each light measured, and a partial ANOVA summary table is shown below.

Source of variation	Sum of squares	Degrees of freedom	Mean square	F
Factor	106,568			X
Error		20		
Total	258,565			

What is the value of the F-statistic labelled as X in the table?

A. 0.21
B. 0.29
C. 3.13
D. 3.51
E. 4.67

10. I have a choice of two routes to get to work. The probability that I choose the first route on any day is 0.6, and the probabilities of my being delayed on the journey are 0.1 for the first route and 0.2 for the second. Calculate the probability of my being delayed exactly once in three days.

A. 0.05
B. 0.31
C. 0.40
D. 0.54
E. 0.86

11. Cars cross a certain point in the highway in accordance with a Poisson process with rate = 3 per minute. If Mr Frogger blindly runs across the highway, what is the probability that he will be uninjured if the amount of time it takes him to cross the road is 5 seconds? (Assume that if he is on the highway when a car passes by, then he will be injured).

A. 0.2212
B. 0.3679
C. 0.6065
D. 0.6321
E. 0.7788

PART B: Please show your work in the space provided.

1. The time, T years, before a machine in a factory breaks down follows the probability density function given by

$$f(t) = \begin{cases} ate^{-bt} & \text{t} > 0 \\ 0 & \text{otherwise} \end{cases}$$

where a and b are positive constants. It may be assumed that, if n is a positive number,

$$\int_0^\infty t^n e^{-bt}\, dt = \frac{n!}{b^{n+1}}.$$

   (a) The mean of T is known to be 1.5. Find the values of a and b.

   (b) Find Var(T).

   (c) Calculate P($T < 1.5$). State, giving a reason, whether this value indicates that the median of T is smaller than the mean of T or greater than the mean of T.

   (d) Calculate the probability that the mean time before breakdown of a random sample of 50 machines is less than 1.75 years.

2. In order to assess the effect of the water discharged from a power station into the nearby Ottawa River, 6 specimens of water were collected from 5 km upstream of the power station, and 8 specimens were collected from 5 km downstream of the power station. The alkalinity of each specimen was measured, in suitable units, with the following results.

Upstream (x_1)	70.3	71.4	68.5	67.3	76.3	72.1		
Downstream (x_2)	151.2	150.8	146.1	145.7	148.2	148.9	147.1	150.4

The mean alkalinity of the water 5 km upstream is μ_1 and the mean alkalinity of the water 5 km downstream is μ_2.

Stating the necessary conditions, test at the 1% significance level whether $\mu_2 - \mu_1 > 70$.

3. (The solution to this problem posted under lecture notes - ANOVA practice problems)

Breitlings sells men's gold, silver and titanium watchbands. A random sample of each type was obtained (in similar styles) and the weight of each watchband (in grams) was recorded. The data are summarized in the table below:

Band type	n_i	$\bar{x}_{i.}$	s_i
1	7	8.157143	0.8403514
2	8	7.975	1.037511
3	8	6.05	1.164965

Conduct an analysis of variance test to determine whether there is any evidence that the mean weights of any two watchband types are different. Include an ANOVA table. Use $\alpha = 0.01$.

4. Many fast food meals contain high amount of sodium. Some research suggests that the amount of sodium is linearly related to the total calories. A sample of 12 sandwiches from a fast food chain was obtained, and a least-square regression was performed using sodium (in grams) as the response variable, and total calories as the explanatory variable.

Use R's output below to answer the questions below:

$Coefficients:$

	$Estimate$	$Std.Error$	$t\ value$	$Pr(>\lvert t \rvert)$
$(Intercept)$	-120.7330	143.8035	-0.840	0.421
$calories$	2.7583	0.3577	7.712	$1.62e-05***$

(a) Write down the equation of the least square regression line.

(b) To test if there is a significant straight line relationship between total calories and sodium, we can perform a t test. What is the value of the t statistic for this test?

(c) From the output, 1.62 X 10^{-5} is the p-value given by R for a t test. Write down the alternative hypothesis of this test.

(d) Test the hypothesis $H_0 : \beta_1 = 0$ versus $H_a : \beta_1 > 0$ at $\alpha = 0.05$.

(e) Compute a 95% confidence interval for the slope β_1.