Practice Test

Set 1

1. A large corporation manufactures one type of item at planes in Atlanta (A), Boston (B) and Cincinnati (C). Atlanta, with new machinery, produces only 2% defective items, 4% of Boston's items are defective. In Cincinnati, things are even worse - 10% of items made there are defective. Of total production of this item, 40% comes from Atlanta, 40% from Boston, and 20% from Cincinnati. All production is sent to a common warehouse. Suppose we choose one item at random from this warehouse. Compute

(a) the probability that the item is defective.

(b) the conditional probability that the part is manufactured in Atlanta given that it is a defective item.

2. X takes on values $-1, 1, 2, 3, 4$. The probability mass function of X satisfies

$$p(i) = \frac{i}{30}, \qquad \text{for } i = 1, \ldots, 4$$

Find $p(-1)$. Determine the mean and the variance of X.

3. The lifetime of a certain type of electronic device (measured in hours) is a continuous random variable X with probability density function given by

$$f(x) = \begin{cases} a/x^3, & x > 9 \\ 0, & x \leq 9 \end{cases}$$

(a) Find a, then compute the mean and the variance of X.

(b) Find $P(X > 20)$.

(c) What is the probability that of 5 such types of devices exactly 2 will function for at least 20 hours, with the assumption that all devices function independently?

4. Two fair dice are rolled. Assume that X is the largest value obtained on any dice and Y is the sum of the values.

(a) Find the joint probability mass function of X and Y.

(b) Compute $V(X + Y)$.

(c) Are X and Y independent? Explain you answer.

5. Assume that jobs arrive every 10 seconds on average

   (a) What is the variance of the number of jobs arriving per minute?

   (b) What is the probability that there are 3 jobs arriving within the first minute?

   (c) What is the probability of waiting less than or equal to 30 seconds?

6. The daily amount of coffee, in liters, dispensed by a machine located in an airport lobby is a random variable X having a continuous uniform distribution on interval $(7, 10)$.

(a) Find the probability that on a given day the amount of coffee dispensed by this machine will be at most 8.8 liters, and determine its mean and variance.

(b) Given that the amount of coffee dispensed by this machine on a given day does not exceed 8.8 liters, what is the probability that is is greater that 8 liters?

7. Assume that $X \sim N(-1, 3)$ and $Y \sim N(2, 4)$. What is the distribution of $3X + 2Y$?

8. Suppose a student measuring the boiling temperature of a certain liquid observes the readings (in degrees Celsius) 102.5, 101.7, 103.1, 100.9, 100.5, and 102.2 on 6 different samples of the liquid. What is the confidence interval for the population mean at a 95% confidence level?

9. The cure rate for the standard treatment of a disease is 45%. Dr. Smith has perfected a primitive treatment which he claims is much better. As evidence, he says that he has used his new treatment on 50 patients with the disease and cured 25 of them. Is this new treatment better at the 1% significance level? How about at 5% significance level?

10. In a study, the following data were recorded on sorbic acid residuals, in parts per million, in ham immediately after being dipped in a sorbate solution and after 60 days of storage:

Slice	1	2	3	4	5
Before Storage	224	270	400	444	590
After Storage	116	96	239	329	437

(a) Assuming the populations to be normally distributed, is there sufficient evidence, at the 0.01 level of significance, to say that the length of storage influences sorbic acid residual concentrations?

(b) Determine a 95% confidence interval of the mean sorbic acid residuals in ham after 60 days of storage.

11. To find out whether a new serum will arrest leukemia (a kind of cancer), 9 mice, all with an advanced stage of the disease, are selected. Five mice receive the treatment and 4 do not. Survival times, in years, from the time the experiment commenced are as follows:

Treatment	2.1	5.3	1.4	4.6	0.9
No Treatment	1.9	0.5	2.8	3.1	

At the 0.01 level of significance, can the serum be said to be effective in treating the disease? Assume the two populations to be normally distributed.

12. A study was done to study the effect of ambient temperature x on the electric power consumed by a chemical plant y. Other factors were held constant, and the data were collected from an experimental pilot plant.

Ambient temperature (x)	250	285	320	295	265
Chemical plant (y)	27	45	72	58	31

(a) Estimate the slope and intercept in a simple linear regression model. Then predict power consumption for an ambient temperature of 65°F.

(b) Construct a 95% confidence interval for the slope of the population regression line.

13. Listed below are the ages (years) of randomly selected race car drivers (based on data reported in USA Today): 32 32 33 33 41 29 38 32 33 23 27 45 52 29 25

   (a) Construct a 95% confidence interval estimate of the mean age of all race car drivers.

   (b) Use the variance calculated in (a) as the population variance. If we wanted the error in estimating the mean age from the two-sided confidence interval to be 1 year at 95% confidence, what sample size should be used?

14. An experiment was conducted to compare three types of paint for evidence of differences in their wearing qualities. They were exposed to abrasive action and the time in hours until abrasion was noticed was observed. Six specimens were used for each type of paint. The data are as follows.

Type 1	Type 2	Type 3
158 97 282	515 264 544	317 662 213
315 220 115	525 330 525	536 175 614

Determine if the evidence suggests that wearing quality differs for the three paints. If significant differences are found, characterize what they are. Is there one paint that stands out? Use the significance level 5%. Summarize the results in an ANOVA table.

15. The following table lists the monthly sales and advertising expenditures for all of last year by a digital electronics company

Month	Jan	Feb	Mar	Apr	May	Jun
Sales (in 1000s)	100	110	112	115	117	116
Advertising Dollars	5.5	5.8	6	5.9	6.2	6.3
Month	Jul	Aug	Sep	Oct	Nov	Dec
Sales (in 1000s)	118	120	121	120	117	123
Advertising Dollars	6.5	6.6	6.4	6.5	6.7	6.8

(a) Compute the median and the iqr of the monthly sales.

(b) Estimate the linear regression line.

(c) Estimate the mean advertising expenditures if the monthly sales is 105.

16. Assuming that a simple linear regression model is appropriate, find the best predicted systolic blood pressure in the left arm given that the systolic blood pressure in the right arm is 100 mm Hg. Then identify the covariance and correlation of the right arm and the left arm.

Right Arm	102	101	94	79	79
Left Arm	175	169	182	146	144

17. A scientist took a sample from a river. The biochemical oxygen demand (BOD) test is conducted over a period of time in days. The resulting data follow

Time (days):	1	4	6	8	10	14	16	20
BOD (mg/liter)	0.6	1.5	1.9	2.1	2.6	3.7	3.5	3.8

Assuming that a simple linear regression model is appropriate, fit the regression model relating BOD (y) to the time (x). Then determine 95% confidence intervals for the coefficients of the model. Based on the coefficient of determination, describe the relationship between two variables.