Applied Statistics Comprehensive Examination

1. A truck can carry a maximum load of 4000 pounds. A manufacturer wants to ship an order of 50 boxes. The weights of the boxes are normally distributed with mean $\mu = 78$ pounds and standard deviation $\sigma = 12$ pounds. Find the probability that all 50 boxes can be sent in one shipment. If the weights are not normally distributed, will the answer still be approximately correct? Why or why not?

2. Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with mean μ and variance σ^2. Show that $\overline{X}^2$ is a biased estimator of μ^2, and determine the bias. (*Hint:* Find two expressions for Var($\overline{X}$) and equate them.)

3. Suppose that you are to test $H_0 : \mu = 14$ against $H_A : \mu \neq 14$ using a t test and preliminary information suggests that σ is about 1.25. Calculate the required sample size so that the probability is 0.1 of falsely accepting H_0 when the true population mean differs from 14 by 0.5.

4. Determine if

$$F_{X,Y}(x,y) = \begin{cases} 1 - e^{-x-y} & \text{for } x > 0, y > 0 \\ 0 & \text{otherwise} \end{cases}$$

is a distribution function for some random variables X and Y. Give reasons to support your answer.

5. Suppose Y is a geometric random variable with probability density function given by:

$$f(y) = \theta(1-\theta)^{y-1} \text{ for } y = 1,2,\ldots$$

Find the maximum likelihood estimator for θ.

6. For 30 female assistant professors selected at random from Arts and Sciences departments of a large state university, the mean salary for the academic year is \$43,000. The corresponding mean salary for a random sample of 50 male assistant professors is \$43,800. The estimated standard deviations for the two samples are \$1000 and \$1400, respectively. Compute a 95% confidence interval for the difference in mean salary between male and female assistant professors at this university. Give a precise interpretation of this result.

7. In an experiment to study the relationship between hypertension and smoking, the following data were obtained from 180 randomly selected subjects:

	Non-Smoker	Moderate Smoker	Heavy Smoker
Hypertensive	21	36	30
Not Hypertensive	48	26	19

Test an appropriate hypothesis using a 5% level of significance and interpret your result.

Applied Statistics Comprehensive Examination

The accompanying dataset contains the advertised prices of ladies' diamond rings and the carat size of the diamond stones. The rings are made with gold of 20 carats purity and are each mounted with a single diamond stone.

The primary objective of this analysis is to build a statistical model that relates the prices of diamond rings to the weights of their diamond stones. Presumably this model could be used as a guide to pricing diamond rings, for either a dealer setting prices or for a consumer evaluating prices.

Here are some specific issues for you to investigate. Please use graphical displays liberally in your solution. Fit models and use tools for statistical inference when practical given the time constraints of this exam.

1. Identify the response (dependent) and predictor (independent) variables in the context of diamond ring pricing.

2. Display the data in an informative manner relative to your identification of the response and predictor variables, and comment on any patterns you observe.

3. Fit a simple statistical model to the data. List the assumptions of the model you fit, and assess whether the assumptions are satisfied by the data.

4. Interpret the model in the context of the problem. Does your model provide reasonable predictions over a wide range of values of the predictor variable? Are the interpretations of the values of any parameter estimates you found reasonable in the context of the problem?

5. Suggest a transformation or an alternative model that you believe might improve either the quality of the fit or the interpretability of the model. If time permits, implement your suggestion. If you are unable to implement the suggestion in the remaining time, describe how the fit of the model, the satisfaction of the assumptions, or the interpretability would improve.

Diamond Ring Prices and Stone Sizes

Size of Stone (in Carats)	Ring Price (In Singapore Dollars)
.17	355
.16	328
.18	325
.25	642
.19	485
.15	323
.18	462
.28	823
.16	336
.20	498
.23	595
.29	860
.12	223
.26	663
.18	468