Statistics Qualifying Exam

9am-1pm, Wednesday, June 16, 2010

Name :

1. Let X be distributed as chi-square with ν degrees of freedom. Let $Y = aX^2 + bX$, where a and b are constants that do not depend on ν. Find a and b so that Y is an unbiased estimator of ν^2.

2. Consider the order statistics of a random sample of size 8 from a continuous distribution: $Y_1 < Y_2 < \cdots < Y_8$. We will use the interval defined by the next-to-extreme observations, i.e., (Y_2, Y_7) as a confidence interval for the population median. What is the confidence level of this interval?

3. Let X_1 and X_2 have the joint pdf

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} 10x_1x_2^2, & 0 < x_1 < x_2 < 1 \\ 0, & \text{elsewhere} \end{cases}$$

   (a) Find the conditional expectation of X_2 given X_1, $E(X_2|X_1)$.

   (b) Find the pdf of Y_1 where $Y_1 = X_1/X_2$.

4. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Poisson distribution with mean μ.

   (a) Find the limiting distribution of the sample mean, $\overline{X}$.

   (b) Find the limiting distribution of $u(\overline{X}) = \sqrt{\overline{X}}$.

5. Let $X_1, \ldots, X_n$ be a random sample from Poisson distribution with mean θ, i.e.,

$$f(x;\theta) = e^{-\theta}\frac{\theta^x}{x!}\ , x = 0, 1, 2, \ldots; \theta > 0\ .$$

   Let $T = \sum_{i=1}^n X_i$ be a statistic.

   (a) Show that the distribution of T is Poisson distribution with mean $n\theta$.

   (b) Show that T is a complete sufficient statistic for θ.

   (c) Find the conditional distribution of X_1 given $T = t$.

   (d) Find the minimum variance unbiased estimator of θ^2.

6. In a study, the steel rods supplied by two different companies were compared. Ten sample springs were made out of the steel rods supplied by each company and the "bounciness" was studied. The data and its summaries are as follow:

company A (x)	9.3	8.8	6.8	8.7	8.5	6.7	8.0	6.5	9.2	7.0
company B (y)	11.0	9.8	9.9	10.2	10.1	9.7	11.0	11.1	10.2	9.6

$$\overline{x} = 7.95,\ s_x^2 = (1.10)^2\ ;\ \overline{y} = 10.26,\ s_x^2 = (0.57)^2$$

Can you conclude that there is virtually no difference in means between the steel rods supplied by the two companies at the 5% level of significance? Should variances be pooled here?

7. The following data pertain to the demand for a product (in 1000's of units) and its price (in dollars) charged in seven different markets:

Price (x)	11	9	12	10	15	12	6
Demand (y)	145	177	109	135	81	118	218

Note that : $\sum x = 75$, $\sum x^2 = 851$, $\sum y = 983$, $\sum y^2 = 150469$ and $\sum xy = 9785$.

(a) Fill in the ANOVA Table below for the linear regression model:

Source	df	Sum of Squares (SS)	Mean Squares (MS)	F-ratio
Regression				
Error				
Total (corrected for mean)				

(b) Give a 95% Confidence interval for the slope.

8. Consider a two-factor study where two factors may be considered to have random factor levels. Factor A and B have a levels, b levels, respectively. There are n repetitions in each treatment.

(a) Write down the appropriate model and its assumptions (constraints and distributional assumptions) for this model.

(b) Write down the ANOVA table including the explicit form of sum of squares, degrees of freedom (d.f.) and Expected Mean squares (EMS).

(c) To test whether or not each effect is significantly different, state the appropriate hypotheses and the corresponding F statistics.

(d) Find the point estimators of all of the model parameters.

9. A production plant cost-control engineer is responsible for cost reduction. One of the costly items in his plant is the amount of water used by the production facilities each month. He decided to investigate water usage by collecting 17 observations on his plant's water usage and other variables. He had heard about multiple regression, but since he was quite skeptical he added a column of random numbers to his original observations. Use the attached information.

Data code
X_1 = average monthly temperature(F)
X_2 = average of production (M pounds)
X_3 = number of plant operating days in the month
X_4 = number of persons on the monthly plant payroll
X_5 = two-digit random number
Y = is the monthly water usage (gallons)

(a) Find the fitted regression model of $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \epsilon_i$ and complete the ANOVA table on your answer sheet (not on the exam sheet).

(b) Comment on the role of the variable X_5. when $X_1 = 58.8$, $X_2 = 7107$, $X_3 = 21$, $X_4 = 129$, $X_5 = 52$

(c) Test the hypothesis $H_0 : \beta_1 = \beta_3 = \beta_5 = 0$ v.s. $H_1 : \beta_3 = 0$. Use $\alpha = .05$.

(d) Perform a stepwise regression using a $\alpha = .05$ level of significance for entering and staying.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value
Model				
Error		722691		
Corrected Total		3192632		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	$Pr > \|t\|$
Intercept	1	6487.35183	1371.64753	4.73	0.0006
x1	1	14.11699	5.33024	2.65	0.0226
x2	1	0.21378	0.04703	4.55	0.0008
x3	1	-126.99857	49.44390	-2.57	0.0261
x4	1	-22.37849	7.56462	-2.96	0.0130
x5	1	-1.34918	2.38034	-0.57	0.5822

Number in Model	R-Square	C(p)	AIC	MSE	SSE	Variables in Model
1	0.3978	16.2616	201.8103	128164	1922459	x2
1	0.1708	27.2961	207.2500	176495	2647418	x4
1	0.0817	31.6266	208.9853	195462	2931930	x1
1	0.0079	35.2113	210.2988	211163	3167442	x3
1	0.0043	35.3854	210.3601	211926	3178883	x5
2	0.5742	9.6907	197.9183	97097	1359361	x2 x4
2	0.4885	13.8573	201.0373	116650	1633106	x1 x2
2	0.4223	17.0730	203.1054	131741	1844373	x2 x3
2	0.3979	18.2605	203.8097	137313	1922389	x2 x5
2	0.2737	24.2966	206.9981	165640	2318960	x1 x4
2	0.1812	28.7900	209.0351	186726	2614168	x3 x4
2	0.1708	29.2961	209.2500	189101	2647418	x4 x5
2	0.1382	30.8770	209.9042	196520	2751284	x1 x3
2	0.0911	33.1699	210.8104	207280	2901925	x1 x5
2	0.0118	37.0226	212.2321	225360	3155042	x3 x5
3	0.6319	8.8875	197.4434	90399	1175191	x1 x2 x4
3	0.6268	9.1372	197.6791	91662	1191601	x2 x3 x4
3	0.5929	10.7846	199.1570	99987	1299830	x1 x2 x3
3	0.5774	11.5361	199.7908	103785	1349205	x2 x4 x5
3	0.4892	15.8213	203.0126	125441	1630736	x1 x2 x5
3	0.4225	19.0653	205.1008	141836	1843869	x2 x3 x5
3	0.3475	22.7068	207.1747	160239	2083110	x1 x3 x4
3	0.2746	26.2503	208.9757	178147	2315914	x1 x4 x5
3	0.1812	30.7890	211.0347	201085	2614103	x3 x4 x5
3	0.1481	32.3963	211.7080	209208	2719705	x1 x3 x5
4	0.7670	4.3213	191.6673	61983	743798	x1 x2 x3 x4
4	0.6379	10.5974	199.1654	96344	1156134	x1 x2 x4 x5
4	0.6293	11.0144	199.5636	98628	1183531	x2 x3 x4 x5
4	0.5935	12.7516	201.1287	108139	1297663	x1 x2 x3 x5
4	0.3485	24.6612	209.1502	173343	2080114	x1 x3 x4 x5
5	0.7736	6.0000	193.1779	65699	722691	x1 x2 x3 x4 x5