Statistics Qualifying Exam

April 2014

1. Suppose that X has p.d.f. $f_\theta(x) = 2\,(\theta - x)\theta^{-2},\ 0 < x < \theta$, where $\theta\ \in \Re^{+}$ is an unknown parameter. We are given $\alpha \in (0,1)$. Use $U = \quad /\theta$ and derive a $100(1-\alpha)\%$ two-sided confidence interval for θ.

2. Let ${}_1$ and ${}_2$ have joint p.d.f. $f(x_1, x_2) = \ 2\ \ {}^{-x_1-x_2}$, $0 < x_2 < x_1 < \infty$, zero elsewhere. Find the p.d.f. of $Y = \ {}_1 - \ {}_2$.

3. Let ${}_1,\ {}_2, \ldots.,\ {}_n$ be i.i.d. Binomial (1 , θ), where $0 < \ \theta < 1$.
   (a) Find a sufficient statistic for θ.
   (b) Find the UMVUE of θ^2.
   (c) Find the UMVUE of $(\theta + 1)^n$. (*Hint for (c):* Look at the m.g.f. of the statistic you found in (a).)

4. Let ${}_1,\ {}_2, \ldots.,\ {}_n$ be i.i.d. with p.d.f. $f_\theta\ (x) = \ \frac{1}{\theta}\quad {}^{-x/\theta}$, for $x > 0$, and zero elsewhere. Here $\theta > 0$.
   (a) Find the best (most powerful) test for testing $H_0\!: \theta = 2$, against $H_1\!: \theta = 3$, at a given level of significance α. Show that the test is based on a simple statistic Y.
   (b) What is the null distribution of Y, and which well-known tables must be used to carry out this test?

5. Let ${}_1,\ {}_2, \ldots.,\ {}_n$ be i.i.d. with the Gamma ($\alpha = 2$, $\beta = \theta$) distribution, where $\theta > 0$.
   (a) Find the MLE of θ and show it is unbiased.
   (b) Is the MLE you found an efficient estimator of θ? PLEASE SHOW AND JUSTIFY YOUR WORK!

6. The following is part of ANOVA table for a simple linear regression model $Y = \beta_0 + \beta_1 \quad + \epsilon$, where ϵ 's are i.i.d. from $N(0, \sigma^2), i = 1, \ldots, n$, and n is the number of observations.
   (a) Compute the coefficient of determination (R^2) . Interpret the value obtained.
   (b) Assume it is known that the least square estimate of β_1 is b_1 =3.57. Construct a t-test of whether or not $\beta_1 = 3$. State the null and alternative hypotheses, the value of the test statistic, the sampling distribution under the null hypothesis and the decision rule.

```
                              Analysis of Variance
                                    Sum of           Mean
Source                   DF        Squares          Square    F Value    Pr > F
Model                    ***        ******          252378     105.88    <.0001
Error                    23         ******       *********
Corrected Total          ***        *******
```

7. An experiment involves 3 factors (Factor A, B, and C). All factors were at two levels (1: low, 2: high), and the number of replicates for each treatment combination was $n = 3$. The cell means ($\bar{y}_{\ \ k.}$) are as follows:

	Factor C (k=1)		Factor C (k=2)	
	Factor B (j=1)	Factor B (j=2)	Factor B (j=1)	Factor B (j=2)
Factor A (i=1)	$\bar{Y}_{111.} = 36.1$	$\bar{Y}_{121.} = 52.3$	$\bar{Y}_{112.} = 56.5$	$\bar{Y}_{122.} = 71.9$
Factor A (i=2)	$\bar{Y}_{211.} = 46.9$	$\bar{Y}_{221.} = 64.1$	$\bar{Y}_{212.} = 68.3$	$\bar{Y}_{222.} = 83.5$

Assume that the data were analyzed with the following fixed effects model $Y_{\ \ kl} = \mu + \alpha\ + \beta\ + \gamma_k + (\alpha\beta)\ \ + (\alpha\gamma)_{\ k} + (\beta\gamma)_{\ k} + (\alpha\beta\gamma)_{\ \ k} + \epsilon_{\ \ kl}$, the random errors $\epsilon_{\ \ kl} \sim N(0, \sigma^2)$.

(a) Fill in the numbered blanks (1) and (2) in the following ANOVA table. Show your steps.

Source	DF	SS
A	1	(1)
B	1	1539.201667
C	1	2440.166667
AB	1	0.240000
AC	1	0.201667
BC	1	2.940000
ABC	1	(2)
Error	16	53.740000
Corrected Total	23	4825.998333

(b) Given $F_{1,16,0.01} = 8.53$. Discuss the significance of effects based on the above ANOVA table (including interaction effects and main effects when it is meaningful). Use $\alpha = 0.01$ in the hypothesis testing for each effect.

(c) Estimate a 99% confidence interval for $\mu_{2..} - \mu_{1..}$. Interpret it.

8. Let $_1, \ _2, \dots, \ _n$ denote a random sample from a $N(\mu_1, \sigma^2)$ population and $Y_1, Y_2, \dots, Y_m$ denote a random sample from a $N(\mu_2, \sigma^2)$ population, σ^2 is unknown. For testing

$$H_0: \mu_1 = \mu_2 \quad vs. \quad H_1: \mu_1 \neq \mu_2 \ ,$$

the same conclusions should be obtained by the two-sample t-test and the fixed effect model. Clearly construct test procedures by these two methods and show their equivalence.

9. The following data reflect information from 17 U.S. Naval hospitals at various sites around the world. The regressors are workload variables, that is, items that result in the need for personnel in a hospital. A brief description of the variables is as follows.

Y=monthly labor-hours/1000
$_1$ =average daily patient load/100
$_2$ =monthly X-ray exposure/1000
$_3$ =monthly occupied bed-days/1000
$_4$ =eligible population in the area/1000
$_5$ =average length of patient's stay, in days

The goal is to produce an appropriate model that will estimate (or predict) personnel needs for Naval hospitals. Normal linear regression models are fitted to the data.

(a) Based on the SAS output in Table 1 select the “best” model using the stepwise method. Use $\alpha = 0.05$.
(b) Use the partial F-test to compare the following two models. Use $\alpha = 0.05$.
$Y = \beta_0 + \beta_2 \ {}_2 + \epsilon$
$Y = \beta_0 + \beta_1 \ {}_1 + \beta_2 \ {}_2 + \beta_3 \ {}_3 + \epsilon$

Percentage Points for the F-distribution $F_{v_1,v_2,0.05} = F^*$ implies $P(F_{v_1,v_2} > F^*) = 0.05$

$F_{1,12,0.05} = 4.75$; $F_{1,13,0.05} = 4.67$; $F_{1,14,0.05} = 4.60$; $F_{1,15,0.05} = 4.54$; $F_{1,16,0.05} = 4.49$;

$F_{2,12,0.05} = 3.89$; $F_{2,13,0.05} = 3.81$; $F_{2,14,0.05} = 3.74$; $F_{2,15,0.05} = 3.68$; $F_{2,16,0.05} = 3.63$.

Table 1 SAS OUPUT for Problem 9

Number in Model	R-Square	Adjusted R-Square	SSE	Variables in Model
1	0.9722	0.9703	13.76231	x3
1	0.9715	0.9696	14.09985	x1
1	0.8934	0.8862	52.76006	x2
1	0.8843	0.8766	57.25318	x4
1	0.3348	0.2904	329.10538	x5
2	0.9867	0.9848	6.57238	x2 x3
2	0.9861	0.9841	6.86819	x1 x2
2	0.9848	0.9826	7.54209	x3 x5
2	0.9840	0.9817	7.90007	x1 x5
2	0.9754	0.9718	12.19065	x3 x4
2	0.9741	0.9704	12.79853	x1 x4
2	0.9725	0.9686	13.59958	x1 x3
2	0.9306	0.9207	34.33937	x2 x4
2	0.9239	0.9130	37.63959	x2 x5
2	0.9104	0.8976	44.31986	x4 x5
3	0.9901	0.9878	4.91340	x2 x3 x5
3	0.9894	0.9870	5.24179	x1 x2 x5
3	0.9873	0.9844	6.26972	x1 x2 x3
3	0.9868	0.9837	6.55484	x2 x3 x4
3	0.9861	0.9829	6.86734	x1 x2 x4
3	0.9850	0.9816	7.40779	x1 x3 x5
3	0.9850	0.9815	7.42033	x3 x4 x5
3	0.9847	0.9811	7.58999	x1 x4 x5
3	0.9785	0.9735	10.63777	x1 x3 x4
3	0.9523	0.9412	23.61614	x2 x4 x5
4	0.9908	0.9877	4.54592	x2 x3 x4 x5
4	0.9906	0.9875	4.64401	x1 x2 x4 x5
4	0.9905	0.9874	4.67752	x1 x2 x3 x5
4	0.9879	0.9838	5.99362	x1 x2 x3 x4
4	0.9851	0.9801	7.38889	x1 x3 x4 x5
5	0.9908	0.9867	4.53505	x1 x2 x3 x4 x5