EXAM II—-STT862 Theory of Statistics and Probability II

Name (Print):______________________ A#______________________

If you sign below, your score will be posted on the course webpage based on the last four digits of the A#

Signature: ______________________ (**No credit will be given unless you show all of your work**.)

1. A clinical trial was carried out to investigate the effect of the drug stelazine on chronic schizophrenics. Patients were divided into two groups matched for age, length of time in the hospital, and score on a behavior rating sheet. A member of each pair was given stelazine and the other a placebo. The following is the behavioral rating scores for the patients at the beginning of the trial and after 3 months. High scores are good.

```
Stelazine before  1.9  2.3  2.0  1.6  1.6  2.6  1.7
          after   1.45 2.45 1.81 1.72 1.63 2.45 2.18
Placebo   before  1.9  2.4  2.0  1.5  1.5  2.7  1.7
          after   1.91 2.54 1.45 1.45 1.54 1.54 1.54
```

(a) Is this a two independent samples problem? Why? (5%)

(b) Carry out a t test at 5% level and draw a conclusion to see whether stelazine is associated with improvement in the patients' scores. (5%)

(c) Carry out a nonparametric test at 5% level to answer the same question in (b). (5%)

(d) Provide a 95% confidence interval for μ(Stelazine improvement)$-\mu$(Placebo improvement). (5%)

(e) Carry out a nonparametric test at 5% level to see if Placebo can improve the patients's score? (5%)

2. The Castle Bakery Company supplies wrapped Italian bread to a large number of supermarkets in a metropolitan area. An experimental study was made of the effects of heights of the shelf display (Factor A: bottom, middle, top) and the width of the shelf display (Factor B: regular, wide) on sales of this bakery's bread during the experimental period. Twelve supermarkets, similar in terms of sales volume and clientele, were utilized in the study. The six treatments were assigned at random to two stores. Sales are presented in the following table

	Factor B			
	regular		wide	
Factor A				
bottom	47	43	46	40
middle	62	68	67	71
top	41	39	42	46

(a) Provide an ANOVA table. (5%)

(b) Carry out a test for interaction at 5% level and make your conclusion. (5%).

(c) Carry out tests for the main factor A and B effects both at 5% level and make your conclusion. (5%).

(d) Use Tukey all pairwise comparison procedure to test the differences among the shelf height means at overall 5% level and make conclusions. (5%).

(e) Obtain 90% simultaneous confidence intervals for the mean sales of regular display width at middle and top shelf heights. (5%).

3. A laboratory tested tires for tread wear by running the following experiment. Tires of a certain brand were mounted on a car. The tires were rotated from position to position every 1000 miles, and the groove depth was measured in mils (.001 inches) initially and after every 4000 miles. Measurements were made at six equiangular positions on each of six grooves around the circumference of every tire. The following table gives the averages of the six equiangular measurements on the outside groove of one tire after every 4000.

Milage (in 1000 miles)	0	4	8	12	16	20	24	28	32
Groove depth (in mils)	394.33	329.50	291.00	255.17	229.33	204.83	179.00	163.83	150.33

(a) Fit a simple linear regression model to the data and find the $\hat{\beta}_0$ and $\hat{\beta}_1$ of the least squares regression line $\hat{Y} = \beta_0 + \beta_1 X$. (5%).
(b) What is the mean groove depth for all tires driven for $25,000$ miles under the experimental conditions? (5%)
(c) What percent of the variation in tread wear is accounted for by the linear regression on milage? Why? (5%)
(d) Provide a 95% confidence interval for the wear rate $-\beta_1$ and test $H_0 : -\beta_1 \geq 8$ against $H_a : -\beta_1 < 8$ at 5% level. (5%).
(e) Provide a 95% confidence interval for the variance σ^2 of the error terms in the simple linear model. (5%).
(f) Based on the plot of residuals, comment on the goodness of the fit of a linear model. (5%).

4. Consider a linear model with $Y_i = \sum_{j=1}^{p} C_{ij}\beta_j + \epsilon_i$ with ϵ_i $(i = 1, \cdots, n)$ being i.i.d $N(0, \sigma^2)$. Let $Y = (Y_1, \cdots, Y_n)'$, $\beta = (\beta_1, \cdots, \beta_p)'$, $C_j = (C_{1j}, \cdots, C_{nj})'$ and $C = (C_1, \cdots, C_p)$. Assume that $C_1, \cdots, C_p$ are linearly independent. Let $\hat{\beta} = (C'C)^{-1}C'Y$.

(a) Show that $\|Y - C\hat{\beta}\|^2/\sigma^2$ has a χ^2 distribution with $n - p$ degrees of freedom. (10%).
(b) Show that $s^2 = \frac{\|Y - C\hat{\beta}\|^2}{n-p}$ is the UMVUE (uniformly minimum variance unbiased estimator) of σ^2. (10%).