Dilek Güvenç
MATH 260

= 4 =

Name:......................

**Date:** May 18, 2007
**Time:** 9:00 – 11:00
**Instructor:** Dilek Güvenç

# MATH 260 FINAL EXAM

| 1 | 2 | 3 | 4 | 5 | TOTAL |
|---|---|---|---|---|---|
| | | | | | |
| 20 | 20 | 20 | 20 | 20 | 100 |

**Important:**

- Check that there are 5 questions in your booklet.
- Answers without intermediate steps may not get any credit
- Do NOT use your mobile phone as a calculator, turn it off during the exam.
- Write your name on each page
- Please, be neat and well organised in your solutions.

*********** GOOD LUCK & HAVE A NICE SUMMER!***********

---

1. Let $X_1, X_2 \ldots X_n$ be a random sample of random variable $X$, where $f(x)$ is gamma density with parameters $\alpha = 3$ and unknown $\beta$.

**a)** Find maximum likelihood estimator, $\hat{\beta}$ of $\beta$. **( 8 points )**
**b)** Is $\hat{\beta}$ an unbiased estimator for $\beta$? Why? **( 6 points )**
**c)** Find $Var(\hat{\beta})$. **( 6 points)**

$$f(x) = \frac{1}{\Gamma(3)\beta^3} x^2 e^{-\frac{x}{\beta}} \quad x>0 \qquad E(X)=3\beta$$

a) $$L_X(\beta) = \prod_{i=1}^{n} \frac{1}{\Gamma(3)\beta^3} x_i^2 e^{-\frac{x_i}{\beta}}$$

$$K_X(\beta) = \ln L_X(\beta) = \sum_{i=1}^{n} \ln\left(\frac{1}{\underbrace{\Gamma(3)}_{2}\beta^3} x_i^2 e^{-\frac{x_i}{\beta}}\right) = \sum_{i=1}^{n}\left(-\ln 2 - 3\ln\beta + 2\ln x_i - \frac{x_i}{\beta}\right)$$

$$= -n\ln 2 - 3n\ln\beta + 2\sum_{i=1}^{n}\ln x_i - \frac{\Sigma x_i}{\beta}$$

$$\frac{dK_X(\beta)}{d\beta} = -\frac{3n}{\beta} + \frac{\Sigma x_i}{\beta^2} = 0 \implies \Sigma x_i = 3n\beta \quad \text{M.L.E of } \beta,\ \hat{\beta} = \frac{\Sigma X_i}{3n} = \frac{\bar{X}}{3}.$$

b) $E(\hat{\beta}) = E\left(\frac{\bar{X}}{3}\right) = \frac{1}{3}E(\bar{X}) = \frac{1}{\cancel{3}}\cdot\cancel{3}\cdot\beta = \beta$. Yes, because $E(\hat{\beta}) = \beta$.

c) $Var(\hat{\beta}) = Var\left(\frac{\bar{X}}{3}\right) = \frac{1}{9}Var(\bar{X}) = \frac{1}{9}\frac{3\beta^2}{n} = \frac{\beta^2}{3n}$.

NAME:.................

**2.** Researchers at a university have said that while older workers were often more productive than their younger counterparts, supervisors tended to rate the older workers lower. Suppose we have the following data, giving the actual productivity of older workers and the supervisor's rating:

| | WORKERS | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Actual Productivety | 82 | 91 | 50 | 82 | 75 | 89 | 90 | 76 |
| Supervisor's Rating | 77 | 80 | 65 | 80 | 70 | 83 | 79 | 75 |

**a)** Does the above data support the contention that older workers are underrated by the supervisors? Test at $\alpha = 0.05$. State the assumption for your answer to be valid. **(10 points)**

**b)** Construct a 90% confidence interval for the true mean difference of actual productivity and supervisor's ratings. Compare the test result and the interval. **(10 points)**

$d_i$'s (actual – supervisor's) 5, 11, –15, 2, 5, 6, 11, 1.

$\bar{d} = \frac{\sum d_i}{8} = 3.25 \quad s_D^2 = 67.64286 \quad s_D = 8.2245$

a) Assuming normality of the differences.

$H_0: \mu_D = 0 \qquad T_{cal} = \frac{3.25 - 0}{\frac{8.22}{\sqrt{8}}} = \frac{9.19}{8.22} = 1.12$

$H_A: \mu_D > 0$

$t_{0.05}(7) = 1.895$ Since $T_{cal} = 1.12 \not> t_{0.05}(7) = 1.895$ we do not reject $H_0$. Conclude that there is no enough evidence to claim that older workers are underrated by the supervisor

b) $1-\alpha = 0.90 \quad \frac{\alpha}{2} = 0.05 \quad t_{0.05}(7) = 1.895$

90% C.I for the true mean difference

$3.25 \mp (1.895)\frac{8.22}{\sqrt{8}}$

$3.25 \mp 5.51$ or $(-2.26, 8.76)$

Since the interval includes 0, using the interval we conclude that there might be no difference between actual productivity and supervisor's ratings

NAME:.................

3. Does the tendency in worker injuries depend on the length of time that a worker has been on the job? An analysis of 714 worker injuries by one manufacturer gave the results shown in the table for the distribution of injuries over the eight 1-hour time periods per shift.

| Hour of shift | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Number of accidents | 93 | 71 | 79 | 72 | 98 | 89 | 102 | 110 |

a) Do the data imply that the probabilities of worker accidents are higher in some time periods than in others? (10 points)

b) Do the data provide sufficient evidence to indicate that the probability of an accident during the last 4 hours of a shift is greater than during the first 4 hours? Test using $\alpha = 0.10$. (10 points)

a) $H_0: p_1 = p_2 = \cdots = p_8 = \frac{1}{8} = 0.125$

$H_A$: For at least one $i$, $p_i \neq 0.125$. $n = 714$

Under $H_0$ $E(X_i) = (714)(0.125) = 89.25$

$$\chi^2_{cal} = \frac{(93-89.25)^2}{89.25} + \frac{(71-89.25)^2}{89.25} + \frac{(79-89.25)^2}{89.25} + \frac{(72-89.25)^2}{89.25}$$
$$+ \frac{(98-89.25)^2}{89.25} + \frac{(89-89.25)^2}{89.25} + \frac{(102-89.25)^2}{89.25} + \frac{(110-89.25)^2}{89.25}$$

$= 15.9$ Let $\alpha = 0.05$ $\chi^2_{0.05}(7) = 14.07$

Since $\chi^2_{cal} = 15.9 > \chi^2_{0.05}(7) = 14.07$ we reject $H_0$. Conclude that probabilities of worker accidents are not the same for all time periods. (Some are higher)

b)

| | First 4 hours | Last 4 hours |
|---|---|---|
| # of accidents | 315 | 399 |

$H_0: p_L = 0.5$ $(p_F = 0.5)$

$H_A: p_L > 0.5$

$\hat{p}_L = \frac{399}{714} = 0.559$

$$Z_{cal} = \frac{0.559 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{714}}}$$

$= 3.153$ $Z_{0.10} = 1.28$

Since $Z_{cal} = 3.153 > Z_{0.1} = 1.28$ we reject $H_0$

Conclude that probability of accident during the last 4 hours of a shift is greater than during the first 4 hours.

RESERVE

NAME:.................

4. Suppose a fire insurance company wants to relate the amount of fire damage in major residential fires to the distance between the residence and the nearest fire station. The study is to be conducted in a large suburb of a major city; a sample of 15 recent fires in this suburb is selected. The amount of damage $Y$ and distance $x$ between the fire and the nearest fire station are recorded for each fire. Fitted regression line is found to be

$$\hat{y} = 10.278 + 4.919x$$

with $s_{\hat{\beta}_0} = 1.42$ , $s_{\hat{\beta}_1} = 0.39$ , $s^2 = 5.37$ and $R^2 = 0.92$.

**a)** Estimate the increase in expected value of fire damage, per unit mile increase in distance with a 95% confidence interval. **(6 points)**

**b)** Show that

$$Var(\hat{\beta}_0) = \sigma^2(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}).$$

**(7 points)**

**c)** Predict the fire damage $y$ if the distance from fire station, $x$ is 4 miles, with a 95% interval. **(7 points)**

a) C.I for $\beta_1$: $\hat{\beta}_1 \mp t_{0.025}^{(13)} S_{\hat{\beta}_1}$

$4.919 \mp (2.16)(0.39)$

$4.919 \mp 0.8424$ or $(4.0766, 5.7614)$

b) $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$

$Var(\hat{\beta}_0) = Var(\bar{Y}) + Var(\hat{\beta}_1 \bar{X})$ because $Cov(\bar{Y}, \hat{\beta}_1) = 0$

$= \frac{\sigma^2}{n} + \bar{X}^2 Var(\hat{\beta}_1)$

$= \frac{\sigma^2}{n} + \bar{X}^2 \frac{\sigma^2}{S_{XX}}$

$= \sigma^2(\frac{1}{n} + \frac{\bar{X}^2}{S_{XX}})$

c) $x=4$ 95% prediction interval

$10.278 + 4.919(4) \mp t_{0.025}^{(13)} S \sqrt{1 + \frac{1}{n} + \frac{(4-\bar{X}^2)}{S_{XX}}}$

$S_{\hat{\beta}_1}^2 = \frac{S^2}{S_{XX}} \Rightarrow S_{XX} = \frac{S^2}{S_{\hat{\beta}_1}^2} = \frac{5.37}{(0.39)^2} = 35.31$

$(1.42)^2 = (5.37)(\frac{1}{15} + \frac{\bar{X}^2}{35.31})$

$(1.42)^2 = 0.358 + (0.15)\bar{X}^2$

$\Rightarrow \bar{X}^2 = 11.056$ $\bar{X} = 3.33$

The interval

$29.954 \mp (2.16)(2.32)\sqrt{1 + \frac{1}{15} + \frac{49.79}{35.31}}$

$29.954 \mp 7.886$

or $(22.068, 37.84)$

RESERVE-4

NAME:..................

5. PPFPO is a viscous liquid used extensively in the electronics industry as a lubricant. In a study the optical density ($Y$) for the prominent infrared absorption of PPFPO was recorded for 12 different experimental settings of band frequency ($x_1$) and film thickness ($x_2$). MINITAB printout for the suggested model is shown below. Answer the following questions by inspecting the output.
   a) What is the suggested regression model in terms of $x_1$, $x_2$ and Y? **(4 points)**
   b) Is there sufficient evidence (at $\alpha = 0.01$) of interaction between band frequency and film thickness? **(3 points)**
   c) What percentage of variation in optical density can be explained by this model? **(3 points)**
   d) Conduct a test of overall significance of the model. Use $\alpha = 0.001$. **(4 points)**
   e) What is the estimate of the error variance? **(3 points)**
   f) Estimate $\beta_2$ with a 90% confidence interval. **(3 points)**

```
The regression equation is
DENSITY = - 0.214 + 0.000257 FREQ - 3.72 THICK + 0.00497 FREQ_THICK


Predictor         Coef     SE Coef      T      P
Constant       -0.2143      0.6866  -0.31  0.763
FREQ         0.0002567   0.0007157   0.36  0.729
THICK          -3.7200      0.9147  -4.07  0.004
FREQ_THICK   0.0049655   0.0009535   5.21  0.001


S = 0.205676   R-Sq = 96.0%   R-Sq(adj) = 94.5%


Analysis of Variance

Source          DF      SS      MS      F      P
Regression       3  8.0481  2.6827  63.42  0.000
Residual Error   8  0.3384  0.0423
Total           11  8.3865
```

a) $E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$

b) P-Value for $\beta_3$ is $0{,}001 < 0{,}01$, yes there is sufficient of interaction betwen $X_1$ and $X_2$.

c) $R^2 = 0{,}96$

d) $H_0: \beta_1 = \beta_2 = \beta_3 = 0$

$H_A: \beta_1 \neq 0$ or $\beta_2 \neq 0$ or $\beta_3 \neq 0$

P value of F test $\sim 0$. $\Rightarrow$ model is overall significant

e) $S^2 = 0{,}0423 = MSE$

f) $-3{,}72 \mp t_{0{,}05}(8)(0{,}9147)$ or $-3{,}72 \mp (1{,}86)(0{,}9147)$

$-3{,}72 \mp 1{,}7 \rightarrow (-5{,}42 \; -2{,}02)$