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

Dept:.................

**Date:** May 23, 2015
**Time:** 15:30-17:40
**Instructor:** Dilek Güvenç

# MATH 260 FINAL EXAM

**IMPORTANT**

- Check that there are 4 questions in your booklet.
- Show all your work. Correct results without sufficient explanation and correct notation may not get full credit. Please, be neat.
- Write your name on each page of your booklet.
- Do NOT use your mobile phone as a calculator, turn it off during the exam.
- Do NOT forget to return the tables and the "cheat-sheet", as well.

| 1 | 2 | 3 | 4 | TOTAL |
|---|---|---|---|---|
| | | | | |
| 25 | 25 | 25 | 25 | 100 |

********* **GOOD LUCK!** *********

---

**1.** Let $x_1, x_2, \cdots x_n$ and $y_1, y_2, \cdots y_m$ denote independent observed random samples of size $n$ and $m$ from a normal distributions with the means $\mu_1 = 1$ and $\mu_2 = 2$ and common variance $\sigma^2$.

**a)** Find maximum likelihood estimator for common variance $\sigma^2$. **(18 Points )**

**b)** Is the estimator in part (a) unbiased estimator for $\sigma^2$? Why? **(7 Points)**

a) $X_i \sim N(\mu_1=1,\sigma^2)$ $\quad Y_j \sim N(\mu_2=2,\sigma^2)$ Likelihood Function

$$L(\sigma^2) = \left(\prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(X_i-1)^2}{\sigma^2}}\right)\cdot\left(\prod_{j=1}^{m}\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(Y_j-2)^2}{\sigma^2}}\right) \text{ (independent samples)}$$

$$K(\sigma^2) = \ln L(\sigma^2) = \sum_{i=1}^{n} \ln\left[\frac{1}{\sqrt{\sigma^2}\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(X_i-1)^2}{\sigma^2}}\right] + \sum_{j=1}^{m} \ln\left[\frac{1}{\sqrt{\sigma^2}\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(Y_j-2)^2}{\sigma^2}}\right]$$

$$= \sum_{i=1}^{n}\left[-\frac{1}{2}\ln\sigma^2 - \ln\sqrt{2\pi} - \frac{1}{2}\frac{(X_i-1)^2}{\sigma^2}\right] + \sum_{j=1}^{m}\left[-\frac{1}{2}\ln\sigma^2 - \ln\sqrt{2\pi} - \frac{1}{2}\frac{(Y_j-2)^2}{\sigma^2}\right]$$

$$= -\frac{n}{2}\ln\sigma^2 - n\ln\sqrt{2\pi} - \frac{1}{2\sigma^2}\sum(X_i-1)^2 - \frac{m}{2}\ln\sigma^2 - m\ln\sqrt{2\pi} - \frac{1}{2\sigma^2}\sum(Y_j-2)^2$$

$$\frac{dK(\sigma^2)}{d\sigma^2}=0 \Rightarrow -\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\sum(X_i-1)^2 - \frac{m}{2\sigma^2}+\frac{1}{2\sigma^4}\sum(Y_j-2)^2 = 0$$

$$\Rightarrow -n\sigma^2 + \sum_{i=1}^{n}(X_i-1)^2 - m\sigma^2 + \sum_{j=1}^{m}(Y_j-2)^2 = 0$$

$\Rightarrow$ MLE of $\sigma^2$, $\quad \hat{\sigma}^2 = \dfrac{\sum(X_i-1)^2+\sum(Y_j-2)^2}{n+m}$

b) $$E(\hat{\sigma}^2) = \frac{1}{n+m}E\left[\sum(X_i-1)^2+\sum(Y_j-2)^2\right] = \frac{1}{n+m}\left[\sum_{i=1}^{n}E(X_i-1)^2+\sum_{j=1}^{m}E(Y_j-2)^2\right]$$

$$= \frac{1}{n+m}\left[\sum_{i=1}^{n}\sigma^2+\sum_{j=1}^{m}\sigma^2\right] = \frac{1}{n+m}(n\sigma^2+m\sigma^2) = \frac{1}{n+m}(n+m)\sigma^2 = \sigma^2$$

$\Rightarrow \hat{\sigma}^2$ is an unbiased estimator

**Name:..................**

**2.** Two different analytical tests can be used to determine the impurity level in steel alloys. Six specimens are tested using both procedures and results are shown in the following table:

| Specimen | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Test 1 | 1.2 | 1.3 | 1.5 | 1.4 | 1.7 | 1.8 |
| Test 2 | 1.4 | 1.7 | 1.5 | 1.3 | 2 | 2.1 |

**a)** Is there sufficient evidence to conclude that tests differ in the mean impurity level, using $\alpha = 0.01$? **(10 Points )**

**b)** Is there evidence to support the claim that Test 1 generates a mean difference 0.1 unit lower than Test 2? Use $\alpha = 0.05$. **( 7 Points )**

**c)** Find 98% confidence interval for the true difference of the means in Test 1 and Test 2 procedures. Compare the interval and the test result in part (a). **( 8 Points )**

$d_i'$s : $-0,2$ , $-0,4$ , $0$, $0,1$ , $-0,3$ , $-0,3$ $\quad n=6$

$\bar{d} = \frac{\sum d_i}{d} = \frac{-1,1}{6} = -0.183 \qquad S_d^2 = \frac{\sum d_i^2 - \frac{(\sum d_i)^2}{6}}{5} = \frac{0.1883}{5} = 0,0376$

$\Rightarrow S_d = 0,194$

a) $H_0: \mu_d = 0$ $\quad$ Test stat. under $H_0$: $T_{cal} = \frac{-0,183}{\frac{0,194}{\sqrt{6}}} = -2,32$

$H_A: \mu_d < 0$

$t_{0,01}(5) = 3,365$ $\quad$ Since $T_{cal} = -2,32 \nless -t_{0,01}(5) = -3,365$ we do not reject $H_0$, conclude that tests do not differ in the mean impurity level at $\alpha = 0,01$.

b) $H_0: \mu_1 - \mu_2 = \mu_d = -0,1$ $\quad$ Test stat. under $H_0$ $T_{cal} = \frac{-0,183+0,1}{\frac{0.194}{\sqrt{6}}} = -1,048$

$H_A: \mu_1 - \mu_2 = \mu_d < -0,1$

$t_{0,05}(5) = 2,015$ $\quad$ Since $T_{cal} = -1,048 \nless -t_{0.05}(5) = -2,015$ we do not reject $H_0$, conclude that $\mu_1 - \mu_2 = -0,1$ at $\alpha = 0.05$

c) 98% C.I for true mean difference

$-0,183 \pm (3.365)\frac{0.194}{\sqrt{6}}$ ; $-0.183 \pm 0.267$ or $(-0,45, 0,084)$

Interval for the true mean difference includes "0". In the test we concluded that there is no difference between mean impurity levels for both tests, so interval supports test result.

——— o ———

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

**3. a)** Show that in simple regression to test the significance of the model, that is; to test $H_0 : \beta_1 = 0$ versus $H_A : \beta_1 \neq 0$ equivalent test statistic is

$$F = \frac{R^2(n-2)}{1-R^2}.$$

**(12 points)**

**b)** What is the probability distribution and parameters of this test statistic? **(3 Points)**

**c)** Suppose that a simple linear regression model has been fit to n= 25 observations. Calculations are : $S_{xx} = 34.789$ , $S_{xY} = 171.25$ and $S_{YY} = 911.517$. Use the test statistic in part (a) to test the significance of the model at $\alpha = 0.10$. **(10 Points)**

a) $H_0: \beta_1 = 0$    These hypotheses test overall significance of
$H_A: \beta_1 \neq 0$    simple regression.

Test statistic $\quad F = \dfrac{SSR/1}{SSE/_{n-2}} \sim F(1, n-2)$

$$F = (n-2)\frac{SSR}{SST-SSR} = (n-2)\frac{SSR/SST}{1-\frac{SSR}{SST}} = \frac{(n-2)R^2}{1-R^2}$$

b) $F \sim F(1, n-2)$

c) $$R^2 = r^2 = \frac{S_{XY}^2}{S_{XX}S_{YY}} = \frac{(171.25)^2}{(34.789)(911.517)} = \frac{29326.5625}{31710.76491} = 0.925$$

$$F = \frac{(0.925)(23)}{1-0.925} = \frac{21.275}{0.075} = 283.67$$

P-Value for this $F \sim 0$, so the model is overall significant.

——— o ———

**Name:.....................**

**4.** Consider the MINITAB output below.

The regression equation is
Y = 517 + 11.5 x1 − 8.14 x2 + 10.9 x3

| Predictor | Coef | SE Coef | T | P |
|---|---|---|---|---|
| Constant | 517.46 | 11.76 | (2) ? | (5) ? |
| x1 | 11.4720 | (1) ? | 36.50 | (6) ? |
| x2 | −8.1378 | 0.1969 | (3) ? | (7) ? |
| x3 | 10.8565 | 0.6652 | (4) ? | (8) ? |

S = 10.2560   R–Sq = ?(9)   R–Sq (adj) = ? (10)

Analysis of Variance

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | (11) ? | 347300 | 115767 | (13) ? | (14) ? |
| Residual Error | 16 | (12) ? | 105 | | |
| Total | 19 | 348983 | | | |

**a)** Fill in the missing values. (Write approximate values of P-values) **( 15 Points)**

**b)** Is the overall model significant at α= 0.01? **(3 Points)**

**c)** Estimate true $\beta_2$ with a 98% confidence interval. **(5 Points)**

**d)** What is the proportion of variability in response explained by this regression model?

**(2 Points)**

a) (1) Since $T = 36.50 = \frac{11.472}{S_{\hat\beta_1}} \Rightarrow S_{\hat\beta_1} = \frac{11.472}{36.5} = 0.3143$

(2) $T = \frac{\hat\beta_0}{S_{\hat\beta_0}} = \frac{517.46}{11.76} = 44.002$ ; (3) $T = \frac{\hat\beta_2}{S_{\hat\beta_2}} = \frac{-8.472}{0.1969} = -41.3296$

(4) $T = \frac{\hat\beta_3}{S_{\hat\beta_3}} = \frac{10.8565}{0.6652} = 16.321$ ; (5) $t_{0.005}(16) = 2.91$ , $T = 44.002 \Rightarrow$ P-Value $\sim 0$

(6) $T = 36.5 \Rightarrow$ P-Value $\sim 0$ ; (7) $T = -41.329 \Rightarrow$ P-Value $\sim 0$

(8) $T = 16.321 \Rightarrow$ P-Value $\sim 0$ ; (9) $R^2 = \frac{SSR}{SST} = \frac{347300}{348983} = 0.995$

(10) $R^2_{ad} = 1 - \frac{20-1}{20-3-1}(1-0.995) = 1 - \frac{19}{16}(0.005) = 0.994$ ; (11) d.o.f = 3

(12) $SSE = SST - SSR = 348983 - 347300 = 1683$ ; (13) $F = \frac{MSR}{MSE} = \frac{115767}{105} = 1102.543$

(14) $F_{0.01}(3,16) = 8.52$   $F = 1102.543 \Rightarrow$ P-Value $\sim 0$

b) Yes. P-Value of $F < 0.01$

c) 98 % C.I for $\beta_2$, $-8.1378 \pm t_{0.01}(16)\, 0.1969$

or $-8.1378 \pm 2.583(0.1969)$

$-8.1378 \pm 0.509$

d) $R^2 = 0.995$