

# Stat312: Final Exam

Moo K. Chung  
mchung@stat.wisc.edu

August 5, 2004

Answer all questions clearly and **circle your final answer**.  
Your answers should be correct up to the **second decimal places**.  
One page note and a calculator are allowed.  
No textbooks, scrap papers or hand-held computers are allowed.  
All variables that are introduced should be defined.  
This exam booklet consists of **five problems** and **eleven pages**.

**Name:** \_\_\_\_\_  
**Student ID:** \_\_\_\_\_

Pledge: On my honor, I have neither given nor received unauthorized aid on this examination.

**Signature:** \_\_\_\_\_

| Question | Points |
|----------|--------|
| 1        |        |
| 2        |        |
| 3        |        |
| 4        |        |
| 5        |        |
| Total    |        |

Please use the following R output in answering the appropriate exam problems.

```
> qnorm(c(0.99,0.975,0.95,0.925,0.9))
[1] 2.33, 1.96, 1.64, 1.44, 1.28
> qt(c(0.99,0.975,0.95,0.925,0.9),2)
[1] 6.96 4.303 2.92 2.28 1.89
> qt(c(0.99,0.975,0.95,0.925,0.9),3)
[1] 4.54 3.18 2.35 1.92 1.64
> qt(c(0.99,0.975,0.95,0.925,0.9),4)
[1] 3.75 2.77 2.13 1.78 1.53
> qchisq(c(0.99,0.975,0.95,0.925,0.9),2)
[1] 9.21 7.38 5.99 5.18 4.60
> qchisq(c(0.99,0.975,0.95,0.925,0.9),3)
[1] 11.34 9.35 7.81 6.90 6.25
> qchisq(c(0.99,0.975,0.95,0.925,0.9),4)
[1] 13.28 11.14 9.49 8.50 7.78
```

1. We wish to fit paired sample  $(x_1, y_1), \dots, (x_n, y_n)$  with a linear regression model  $Y = cx + \epsilon$ . We assume that  $\epsilon$  follows a normal distribution with zero mean and variance  $\sigma^2$ .
  - (a) Estimate  $c$  by minimizing the sum of the squared residuals (5pts).
  - (b) Write the log-likelihood function as a function of  $c$  and  $\sigma$  (5pts).

- (c) Find the likelihood estimator of  $c$  by differentiating the log-likelihood function in (b). Derive everything (5pts).
- (d) Either prove or disprove unbiasedness of the estimator you computed in (c) (5pts, no point given if (c) is incorrect).

- (e) Compute the variance of the estimator you computed in (a) (5pts, no point given if (a) is incorrect).
- (f) If the sample correlation coefficient of the above paired data is 0.5, what is the sample correlation coefficient of data  $(x_1 - \bar{x}, y_1 - \bar{y}), \dots, (x_n - \bar{x}, y_n - \bar{y})$ ?  $\bar{x}$  and  $\bar{y}$  are the respective sample means of  $x_i$ 's and  $y_i$ 's. Prove your result (5pts).

2. When you throw a coin 10 times, you observed 3 heads. Test if the coin is biased at  $\alpha = 0.2$ . Clearly state appropriate parameters, hypotheses, a test statistic. You may use the following R output (15pts).

```
> pbinom(0:5,10,0.5)
[1] 0.000 0.010 0.060 0.170 0.380 0.620
```

3. A manufacturer of automatic washers offers a particular model in one of three colors: white, green and blue. Of the 100 washers sold, 40 were white, 35 green, 25 blue. In each of the subsequent questions, clearly state appropriate parameters, hypotheses, a test statistic and its distribution under the null assumption.
- (a) Would you conclude that customers have no preference? Test it at  $\alpha = 0.05$  (10 pts).

- (b) Would you conclude that customers have a preference for white color? Test it at  $\alpha = 0.05$  (10 pts).

4. A quality control engineer has measured the numbers of defectives per day from a certain production process for 50 days and recorded below. Test if the number of defectives follows a binomial distribution at  $\alpha = 0.05$ . (20pts).

| number of defects | frequencies |
|-------------------|-------------|
| 0                 | 10          |
| 1                 | 24          |
| 2                 | 10          |
| 3                 | 6           |

5. Problem on the coefficient of determination.

- (a) The following is the R output of a regression analysis based on linear model  $Y = \beta_0 + \beta_1 x + \epsilon$  for 11 paired data  $(x_i, y_i)$ . Compute the coefficient of determination and the coefficient of correlation (5pts).

```
>summary(lm(y~x))
Call:
lm(formula = y ~ x)
Residuals:
    Min       1Q   Median       3Q      Max
   -12    -6.6    -0.28     7.64    10.06
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)        -14      16.1   -0.889    0.4
x                   7       1     3      0.001 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- (b) It is shown during the lecture that the coefficient of determination is always between 0 and 1. Prove it (15pts).