

STATS 100C Midterm Exam 2 Information (Spring 2009)

- **Midterm exam 2: Monday, May 18, 2-2:50pm in MS 5128.**
- This is a closed-book exam, but you can use a calculator and the formulas provided.
- The z, t and F tables will be provided if necessary.
- Material: Chapters 1-4, HW 1-6.

Formulas for STATS 100C Midterm Exam 2 (Spring 2009)

Chapter 2: Simple Linear Regression

The pdf of $N(\mu, \sigma^2)$ is $f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp[-\frac{(x-\mu)^2}{2\sigma^2}]$.

Model: $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$. LSE: $\hat{\beta}_1 = \frac{s_{xy}}{s_{xx}}$, $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$, $\hat{\mu}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0$.

Variance: $V(\hat{\beta}_1) = \sigma^2/s_{xx}$, $V(\hat{\beta}_0) = \sigma^2(\frac{1}{n} + \frac{\bar{x}^2}{s_{xx}})$, $V(\hat{\mu}_0) = \sigma^2(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{s_{xx}})$.

Notation: $r = s_{xy}/\sqrt{s_{xx}s_{yy}}$, $s_{xx} = \sum(x_i - \bar{x})^2 = \sum x_i^2 - \frac{1}{n}(\sum x_i)^2$

$s_{xy} = \sum(x_i - \bar{x})(y_i - \bar{y}) = \sum(x_i - \bar{x})y_i = \sum x_i y_i - \frac{1}{n}(\sum x_i)(\sum y_i)$

$SST = \sum(y_i - \bar{y})^2$, $SSR = \sum(\hat{\mu}_i - \bar{y})^2 = \hat{\beta}_1^2 s_{xx}$ and $SSE = \sum(y_i - \hat{\mu}_i)^2$.

Chapter 3: Random Vectors Suppose $E(y) = \mu$ and $V(y) = \Sigma$.

$$E(Ay + b) = A\mu + b, V(Ay + b) = A\Sigma A', V(y) = E[(y - \mu)(y - \mu)'], cov(a'y, b'y) = a'\Sigma b.$$

The pdf of a n -variate normal distribution $N(\mu, \Sigma)$ is

$$f(x) = (2\pi)^{-\frac{n}{2}} |\Sigma|^{-\frac{1}{2}} \exp[-\frac{1}{2}(x - \mu)'\Sigma^{-1}(x - \mu)]$$

Chapter 4: Multiple Linear Regression Model

Model: $y = X\beta + \epsilon$, $\epsilon \sim N(0, \sigma^2 I)$

LSE $\hat{\beta} = (X'X)^{-1}X'y$, $s^2 = S(\hat{\beta})/(n - p - 1) = (y - X\hat{\beta})'(y - X\hat{\beta})/(n - p - 1)$.

Fitted values $\hat{\mu} = X\hat{\beta} = Hy$; residuals $e = y - \hat{\mu} = y - Hy = (I - H)y$; where $H = X(X'X)^{-1}X'$.

Variance $V(\hat{\beta}) = \sigma^2(X'X)^{-1}$, $V(\hat{\mu}) = \sigma^2 H$, $V(e) = \sigma^2(I - H)$, $V(a'\hat{\beta}) = \sigma^2 a'(X'X)^{-1}a$.

Properties: $\hat{\beta}$ and e are independent; $S(\hat{\beta})/\sigma^2 = e'e/\sigma^2 \sim \chi_{n-p-1}^2$; LSE is BLUE.

CI for β_i : $\hat{\beta}_i \pm t(1 - \alpha/2, n - p - 1)s\sqrt{v_{ii}}$; CI for $a'\beta$: $a'\hat{\beta} \pm t(1 - \alpha/2, n - p - 1)s\sqrt{a'(X'X)^{-1}a}$

F statistic for testing $H_0 : A\beta = 0$:

$$F = \frac{\|\hat{\mu} - \hat{\mu}_A\|^2/l}{S(\hat{\beta})/(n - p - 1)} = \frac{(S(\hat{\beta}_A) - S(\hat{\beta}))/l}{S(\hat{\beta})/(n - p - 1)}$$

Joint confidence region for β : $(\hat{\beta} - \beta)'(X'X)(\hat{\beta} - \beta)/[(p + 1)s^2] \leq F(1 - \alpha, p + 1, n - p - 1)$.

$SST = \sum(y_i - \bar{y})^2$, $SSE = \sum(y_i - \hat{\mu}_i)^2$, $SSR = \sum(\hat{\mu}_i - \bar{y})^2 = \hat{\beta}'X'y - n\bar{y}^2 = \hat{\beta}'X'X\hat{\beta} - n\bar{y}^2$.

$R^2 = SSR/SST = 1 - SSE/SST$