STA302/STA1001, Weeks 8-9

Mark Ebden, 31 October & 2 November 2017

With grateful acknowledgment to Alison Gibbs

Plan for Tuesday 31 October

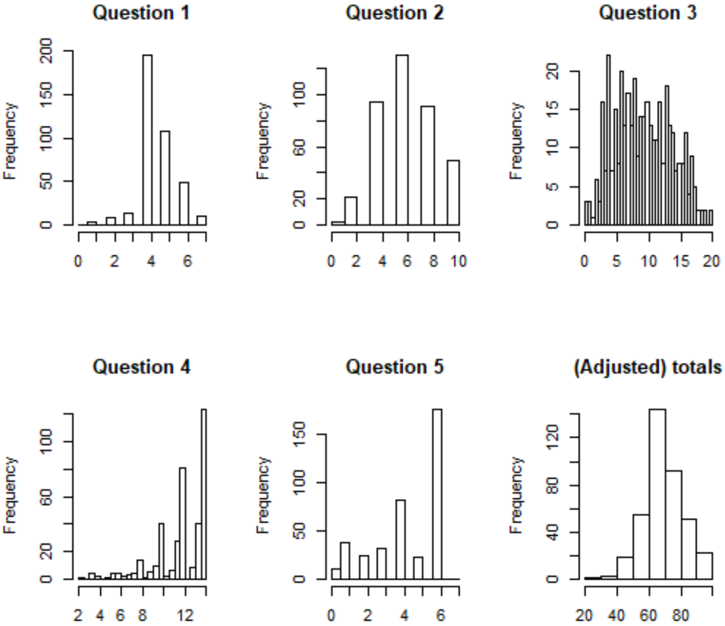
- Section 1 can pick up midterms and digest them for a few minutes
- Midterm discussion
- Chapter 5
- One-to-one discussion about any midterm issues

Midterms: Section 1

The number of test takers was 406 in Section 1: STA302/1001 LEC0101 & LEC2001.

Marking has been completed for Section 1 only.

Midterm marks: Based on 96% of test papers in Section 1



Midterm marks: Section 1

Raw scores:

- Top students got 55/57
- Average of 64%

Adjusted scores:

- Two marks of 100%
- Average of 68.3%

Questions:

- Questions 1, 2, and 5 had averages of 62 to 64%
- Hardest question was #3: average was 47%
- Easiest question was #4: average was 84%
- We'll review some of these in a moment

To request a re-grade: Section 1

By 9 November, please email `sta302sec1@gmail.com` with a description of the problem, including whether or not you spoke with me during class on 31 October.

You should receive a reply within a week of sending your request, and your mark may go up or down.

For efficiency, you may wish to include a picture of the problem. This is encouraged but optional. If the picture you take isn't found to match what we have on record, when verified at a later date, then you may be subject to an academic offence and any previous mark adjustment would be moot.

Post-midterm work so far

- §3.3 (Transformations) except for Box-Cox transformations and inverse-response plots
- §5.2 (Estimation and Inference in MLR) is what we've been heading towards, via the RMA

A closer look at $\mathbf{X}'\mathbf{X}$ and $(\mathbf{X}'\mathbf{X})^{-1}$

Previously we simplified $\mathbf{X}'\mathbf{X}$, for $\mathbf{X} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$

We began to consider $(\mathbf{X}'\mathbf{X})^{-1}$ (Week 8, slide 12), because of its appearance in $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$.

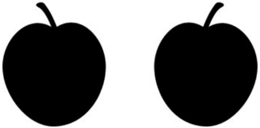
This $\hat{\beta}$ expression is a concise way to write the estimators for linear regression, compared to $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$ and $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}$.

Properties of least squares estimates

Recall that $\mathrm{E}(\hat{\beta}_0) = \beta_0$ and $\mathrm{E}(\hat{\beta}_1) = \beta_1$, and that

$$\mathrm{var}\left(\hat{\beta}_1\right) = \frac{\sigma^2}{S_{xx}}, \qquad \mathrm{var}(\hat{\beta}_0) = \sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right] \quad \text{and} \quad \mathrm{cov}(\hat{\beta}_0, \hat{\beta}_1) = -\frac{\sigma^2 \bar{x}}{S_{xx}}$$

Let's now confirm that our new equations give this as well.



Next steps

- HW2 (not for credit) is on Portal
- We'll continue in Chapter 5
  - More parallels between the old- and new expressions
  - MLR (multiple linear regression)

Appendix: Content for 2 November

Missed lectures this week?

You can still pick up your test before the break. Come to SS 6027 CLTA on Friday 3 November from 10-11 am. If you're unable to attend that time as well, you can email me and send a friend/classmate on your behalf.

This is your last opportunity prior to the deadline for regrading requests (9 November for Section 1).

Recall our R code for β_0 and β_1

e.g. from Weeks 5-6, slide 3, handling question 1 of Chapter 2:

```
X <- read.csv("playbill.csv")
y <- X$CurrentWeek; x <- X$LastWeek
my <- mean(y); mx <- mean(x); n <- length(x)
Sxy <- sum((x-mx)*(y-my)); Sxx <- sum((x-mx)^2)
b1 <- Sxy/Sxx # (2.4), beta-hat-1
b0 <- my - b1*mx # (2.3), beta-hat-0
yHat <- b1*x + b0 # (2.1)
```

Application of the matrix approach to a small but real dataset

```
Q <- read.csv("playbill.csv")
Y <- Q$CurrentWeek; n <- length(Y)
X <- matrix(c(rep(1,n),Q$LastWeek),ncol=2,byrow=FALSE)
BetaHat <- solve(t(X)%*%X)%*%t(X)%*%Y
Yhat <- X%*%BetaHat
print(BetaHat)
```

```
##                [,1]
## [1,] 6804.8860355
## [2,]    0.9820815
```

%*% matrix multiplication

solve() find inverse

Optional material: In “Octave”, `BetaHat = inv(X'*X)*X'*Y`

Fitted values ($\widehat{\mathbf{Y}}$) in matrix form

Recall from slide 21 in Weeks 6–7 that our model is:

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{e}$$

Recalling that $\hat{\beta}$ is unbiased and that $\mathrm{E}(\mathbf{e}) = \mathbf{0}$, we have $\widehat{\mathbf{Y}} = \mathbf{X}\hat{\beta}$. So:

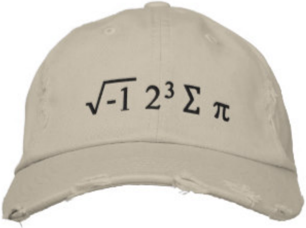
$$\widehat{\mathbf{Y}} = \mathbf{X}\hat{\beta} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \mathbf{H}\mathbf{Y}$$

where $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ is the **hat matrix**, comprised of the h_{ij} values.

Re-"cap"

Recall from weeks 4-5 that h in $h_{ij} = \frac{1}{n} + \frac{(x_i - \bar{x})(x_j - \bar{x})}{S_{xx}}$ stands for "hat". This is because, considering $\hat{y}_i = \sum_{j=1}^{n} h_{ij}\, y_j$, the h values show how to get from y_i's to $\hat{y}_i$'s.

This is even more apparent in the matrix notation.



Properties of $\mathbf{H}$

$\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ is an example of an *idempotent* matrix. **Exercise:** Show this.

$\mathbf{H}$ is symmetric. **Exercise:** Show this.

prove H^2 = H

Five facts about idempotent matrices

1. A square matrix $\mathbf{A}$ is idempotent iff $\mathbf{A}^2 = \mathbf{A}$
2. If $\mathbf{A}$ is idempotent then $\text{trace}(\mathbf{A}) = \text{rank}(\mathbf{A})$
3. $\mathbf{A}$ is idempotent iff $\text{rank}(\mathbf{A}) + \text{rank}(\mathbf{I} - \mathbf{A}) = n$ where the dimensions of $\mathbf{A}$ are $n \times n$ and $\mathbf{I}$ is the $n \times n$ identity matrix
4. For hat matrix $\mathbf{H}$ and matrix of all 1's $\mathbf{J}$, the following matrices are idempotent:

$$\mathbf{H} \qquad \mathbf{I} - \mathbf{H} \qquad \frac{1}{n}\mathbf{J} \qquad \mathbf{H} - \frac{1}{n}\mathbf{J}$$

5. If $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are idempotent and $\mathbf{A} = \mathbf{B} + \mathbf{C}$, then $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{B}) + \text{rank}(\mathbf{C})$

iff = "if and only if"

Residuals ($\widehat{\mathbf{e}}$) in matrix form

The residuals are given by

$$\widehat{\mathbf{e}} = \begin{pmatrix} \hat{e}_1 \\ \vdots \\ \hat{e}_n \end{pmatrix} = \mathbf{Y} - \widehat{\mathbf{Y}} = \mathbf{Y} - \mathbf{X}\hat{\beta}$$

Donning our new hat matrix, this can be rewritten as $\widehat{\mathbf{e}} = \mathbf{Y} - \mathbf{HY}$ before determining $\mathrm{E}(\widehat{\mathbf{e}})$ and var($\widehat{\mathbf{e}}$).

To begin, how could we factorize $\widehat{\mathbf{e}} = \mathbf{Y} - \mathbf{HY}$?

Properties of $\mathbf{I} - \mathbf{H}$

Is $\mathbf{I} - \mathbf{H}$ idempotent?

Is $\mathbf{I} - \mathbf{H}$ symmetric?

Continuing:

$\mathrm{E}(\widehat{\mathbf{e}})=$

$\mathrm{var}(\widehat{\mathbf{e}})=$

We hope you have an enjoyable week

- Remember that for the Study Break of 6-10 November, there will be a pause in lectures, TA office hours, and my office hours
- Around the time of the next lecture, there will be a poll on Piazza to ask how *cumulative* you'd like the exam content to be (four options). The final poll results will be discussed with TAs and a nonzero percentage of pre-midterm material will be set