statweb.stanford.edu/ owen/courses/305

Prerequisites

- Matrix algebra: eigenvalues, rank, Orthogonal matrices...

- Probability: normal, t, χ^2, CLT, covariance

- Statistics: p-value, conf. intervals, hypothesis testing, regression

- Computation: R, python, Matlab, C

- Experience: fitting models, applying methods

Statistics is almot but not quite math
Statistics is almot but not quite computing

Modeling is tricky:

- hard to choose a model

- wrong assumptions can lead to right answers

- can't quite prove things about the world

Linear Models

have X predict $y \in R$
X arbitrary
data $(X_i, y_i) i = 1, ..., n$

Least Square Error criteria:
Best predict of y:
for $X = x, \mu(x) = E(y|X)$

$E([Y - m(x)]^2|X = x) = E([Y - \mu + \mu - m(x)]^2|X = x) = V(Y|X = x) + (\mu - m(x))^2 \geq V(Y)$

However, if you do not believe in the mean square error criteria, the answer is different:
For absolute deviation loss,

$$E(|Y - m(X)||X = x)$$

take $m(X) = median(Y|X = x)$

Alternate Proof (sketch)
Set $\frac{d}{\mathrm{dm}} E((y - m)^2 | X = x) = 0$

Linear Model Examples

$$Y_i = \beta_0 + \beta_1 X_i + e_i$$
$$e \sim (0, \sigma^2)$$

(maybe normal)