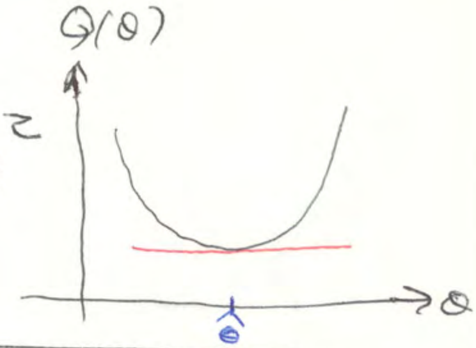
Tues Jan 28

Least squares estimation: Estimate θ with the value that minimizes

$$Q(\theta) = \sum_{i=1}^{n} (X_i - E_\theta(X_i))^2$$



<u>Ex</u> $X_1, \dots X_n \overset{iid}{\sim} Exp(\lambda)$ so $E(X_i) = \frac{1}{\lambda}$

$$Q(\theta) = \sum_{i=1}^{n} (X_i - \tfrac{1}{\lambda})^2$$

$$\frac{dQ}{d\lambda} = \frac{d}{d\lambda} \sum_{i=1}^{n} (X_i - \lambda^{-1})^2 = \sum_{i=1}^{n} \frac{d}{d\lambda} (X_i - \lambda^{-1})^2$$

$$= \sum_{i=1}^{n} 2(X_i - \lambda^{-1})(-1)(-1)\lambda^{-2} = \frac{2}{\lambda^2} \sum_{i=1}^{n} (X_i - \lambda^{-1})$$

$$= \frac{2}{\lambda^2} \left[\sum_{i=1}^{n} X_i - \frac{n}{\lambda} \right] \overset{set}{=} 0 \Rightarrow \sum_{i=1}^{n} X_i = \frac{n}{\lambda}$$

$$\Rightarrow \lambda = \frac{n}{\sum_{i=1}^{n} X_i} = \frac{1}{\bar{X}_n}$$

Second derivative test

$$\frac{d^2Q}{d\lambda^2} = \frac{d}{d\lambda} 2\left(\lambda^{-2} \sum X_i - n\lambda^{-3}\right) = 2\left(\sum X_i (-2)\lambda^{-3} - n(-3)\lambda^{-4}\right)$$

$$= 2\left(-2\frac{\sum X_i}{\lambda^3} + \frac{3n}{\lambda^4}\right) \text{ at } \lambda = \frac{1}{\bar{X}}$$

$$2\left(\frac{3n}{\lambda^4} - 2\,\frac{\Sigma X_i}{\lambda^3}\right) = 2\left(3n\bar{x}^4 - 2n\bar{x}^4\right)$$

$$= 2nx^4 > 0$$ ccup $\cup$ MIN

$$\hat{\lambda} = \frac{1}{\bar{X}_n} = \text{MOM}$$

In general Let $E_\theta(X_i) = g(\theta)$

Minimize

$$Q(\theta) = \sum_{i=1}^{n}(X_i - g(\theta))^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2 + n\left(\bar{X} - g(\theta)\right)^2$$

Pos, SMALL AS possible if $g(\theta) = \bar{X}$

$$\therefore \hat{\theta} = g^{-1}(\bar{X}) = \text{MOM}$$

So why Bother?

$E(X_i)$ might be different

<u>Ex</u> $Y_i = \beta x_i + E_i$

$E_1, \dots, E_n \stackrel{ind}{\sim} ?\,(\mu = 0, \sigma^2)$

$x_1 \dots x_n$ Fixed, known constants

$\theta = (\beta, \sigma^2)$

For ex x_i = drug dose

Y_i = Response to drug

Estimate β by least squares

$$E(Y_i) = E(\beta x_i + E_i) = E(\beta x_i) + E(E_i)$$

$$= \beta x_i + 0 = \beta x_i$$

$$Q(\beta) = \sum_{i=1}^n (Y_i - E_\beta(Y_i))^2 = \sum_{i=1}^n (Y_i - \beta x_i)^2$$

$$\frac{dQ}{d\beta} = \frac{d}{d\beta} \sum_{i=1}^n (Y_i - \beta x_i)^2 = \sum_{i=1}^n \frac{d}{d\beta}(Y_i - \beta x_i)^2$$

$$= \sum_{i=1}^n 2(Y_i - \beta x_i)(-1)x_i = -2\sum_{i=1}^n (x_i Y_i - \beta x_i^2)$$

$$= -2\left[\sum_{i=1}^n x_i Y_i - \beta \sum_{i=1}^n x_i^2\right] \stackrel{set}{=} 0$$

$$\Rightarrow \sum_{i=1}^{n} x_i y_i = \beta \sum_{i=1}^{n} x_i^2 \Rightarrow \beta = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

is it a min

$$\frac{d^2 Q}{d\beta^2} = \frac{d}{d\beta}\left(-2\left(\sum_{i=1}^{n} x_i y_i - \beta \sum x_i^2\right)\right)$$

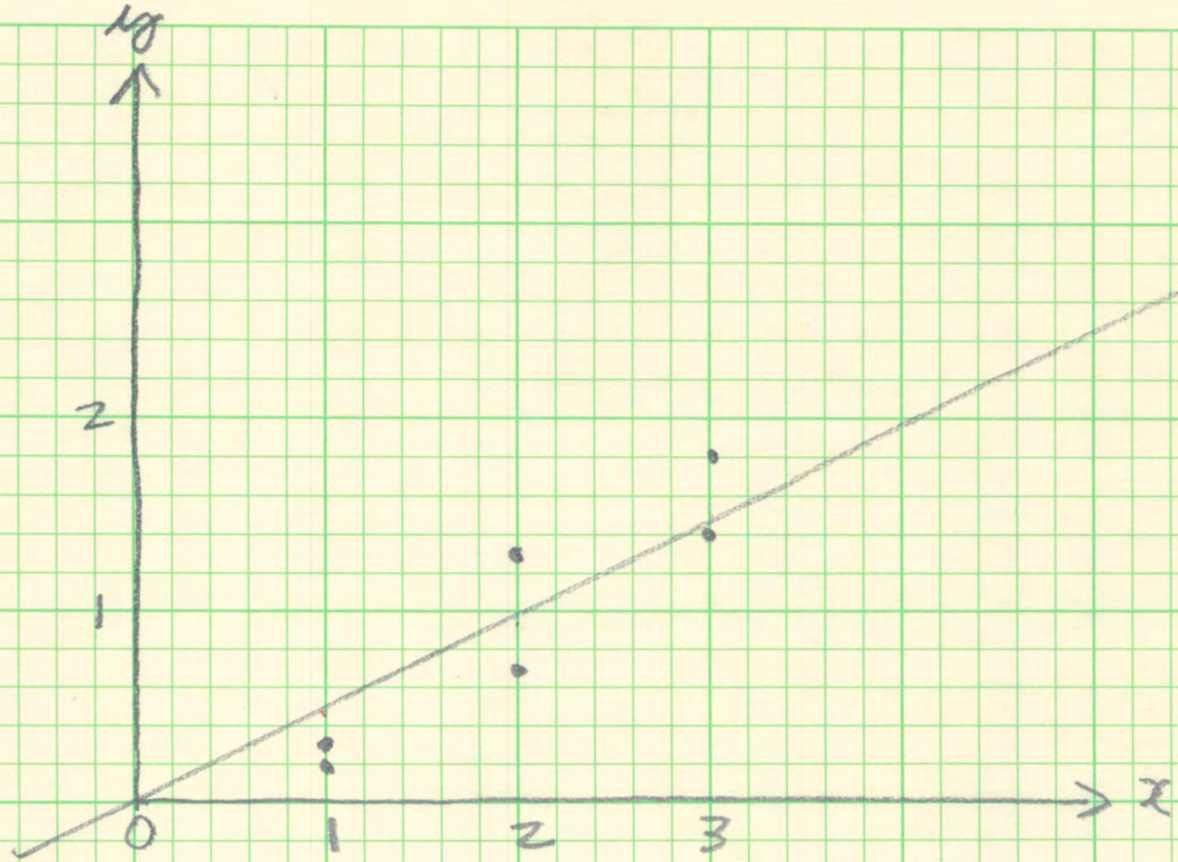
$$= -2\left(-\sum_{i=1}^{n} x_i^2 (1)\right) = 2\sum_{i=1}^{n} x_i^2 > 0$$

CCUP MIN

$$\text{So } \hat{\beta} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

x	y	xy	x^2
1	0.3	0.3	1
1	0.2	0.2	1
2	1.3	2.6	4
2	0.7	1.4	4
3	1.4	~~8.8~~ 4.2	9
3	1.6	~~4.8~~ 4.8	9
12	5.5	13.5	28

$$\hat{\beta} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} = \frac{13.5}{28} = 0.482$$


y
2
1
0
1
2
3
x

IS $\hat{\beta}_n = \dfrac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}$ UNBIASED?

$$E(Y_i) = \beta x_i$$

$$E(\hat{\beta}) = \frac{1}{\sum_{i=1}^{n} x_i^2} E\left(\sum_{i=1}^{n} x_i Y_i\right) = \frac{1}{\sum_{i=1}^{n} x_i^2} \sum_{i=1}^{n} E(x_i Y_i)$$

$$= \frac{1}{\sum_{i=1}^{n} x_i^2} \sum_{i=1}^{n} x_i E(Y_i) = \frac{1}{\sum_{i=1}^{n} x_i^2} \sum_{i=1}^{n} x_i \beta x_i$$

$$= \beta \frac{1}{\cancel{\sum_{i=1}^{n} x_i^2}} \cancel{\sum_{i=1}^{n} x_i^2} = \beta \quad \underline{\text{unbiased}}$$

<u>Consistent?</u> Note $x_1, x_2, \ldots$ is a sequence of fixed constants

Suppose $\dfrac{1}{\sum_{i=1}^{n} x_i^2} \longrightarrow 0$ as $n \to \infty$

Variance rule Need $\lim_{n\to\infty} E(\hat{\beta}_n) = \beta$

$$\lim_{n\to\infty} Var(\hat{\beta}_n) = 0$$

$$\lim_{n\to\infty} Var(\hat{\beta}_n) = \lim_{n\to\infty} Var\left(\frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}\right)$$

$$\lim_{n\to\infty} Var(\hat{\beta}_n) = \lim_{n\to\infty} Var\left(\frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}\right)$$

$$= \lim_{n\to\infty} \frac{1}{\left(\sum_{i=1}^n x_i^2\right)^2} Var\left(\sum_{i=1}^n x_i Y_i\right)$$

Note $Y_i = \beta_0 x_i + E_i$

E_i are independent

$$\overset{ind}{=} \lim_{n\to\infty} \frac{1}{\left(\sum_{i=1}^n x_i^2\right)^2} \sum_{i=1}^n Var(x_i Y_i)$$

$$= \lim_{n\to\infty} \frac{1}{\left(\sum_{i=1}^n x_i^2\right)^2} \sum_{i=1}^n x_i^2 Var(Y_i)$$

$$= \lim_{n\to\infty} \frac{1}{\left(\sum x_i^2\right)^2} \sum_{i=1}^n x_i^2 \sigma^2$$

$$= \lim_{n\to\infty} \frac{\sigma^2 \sum_{i=1}^n x_i^2}{\left(\sum_{i=1}^n x_i^2\right)^2} = \sigma^2 \lim_{n\to\infty} \frac{1}{\sum_{i=1}^n x_i^2}$$

$= 0$ ~~Consistent~~ by

$\hat{\beta}_n$ is consistent by variance rule

$\boxed{}$

Extend the model: $Y_i = \beta_0 + \beta_1 x_i + E_i$

$E_1, \dots E_n \overset{iid}{\sim} (\mu = 0, \sigma^2)$

$x_1, \dots x_n$ fixed observable constants

$$E(Y_i) = E(\beta_0 + \beta_1 x_i + E_i) = \beta_0 + \beta_1 x_i + \underbrace{E(E_i)}_{0}$$

$$= \beta_0 + \beta_1 x_i$$

$$Var(Y_i) = Var(\beta_0 + \beta_1 x_i + E_i) = Var(E_i) = \sigma^2$$

Want Least-squares estimates of β_0 & β_1

$$Q(\beta_0, \beta_1) = \sum_{i=1}^{n} (Y_i - \overbrace{\beta_0 - \beta_1 x_i}^{E(Y_i)})^2$$

$$\frac{\partial Q}{\partial \beta_0} \overset{set}{=} 0 \quad , \quad \frac{\partial Q}{\partial \beta_1} \overset{set}{=} 0$$

Solve 2 EQ in 2 unknowns

$$\frac{\partial Q}{\partial \beta_0} = \frac{\partial}{\partial \beta_0} \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 x_i)^2$$

$$= \sum_{i=1}^{n} \frac{\partial}{\partial \beta_0} (Y_i - \beta_0 - \beta_1 x_i)^2$$

$$= \sum_{i=1}^{n} 2 (Y_i - \beta_0 - \beta_1 x_i)(-1) \overset{set}{=} 0$$

$$\Longrightarrow \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\Longrightarrow \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} \beta_0 - \beta_1 \sum_{i=1}^{n} x_i = 0$$

$$\Longrightarrow \sum_{i=1}^{n} Y_i = n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i$$

$$\Longrightarrow \frac{1}{n} \sum_{i=1}^{n} Y_i = \beta_0 + \beta_1 \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$\Longrightarrow \bar{Y}_n = \beta_0 + \beta_1 \bar{x} \qquad (1)$$

$$\frac{\partial Q}{\partial \beta_1} = \frac{\partial}{\partial \beta} \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 x_i)^2$$

$$= \sum_{i=1}^{n} 2 (Y_i - \beta_0 - \beta_1 x_i)(-x_i) \quad \text{chain rule}$$

$$= -2 \sum_{i=1}^{n} (x_i Y_i - \beta_0 x_i - \beta_1 x_i^2) \overset{set}{=} 0$$

$$\Longrightarrow \sum_{i=1}^{n} x_i Y_i - \beta_0 \sum_{i=1}^{n} x_i - \beta_1 \sum_{i=1}^{n} x_i^2 = 0$$

$$\Rightarrow \sum_{i=1}^{n} x_i Y_i = \beta_0 \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} x_i^2 \qquad (2)$$

and $\bar{Y}_n = \beta_0 + \beta_1 \bar{x}$ $\qquad (1)$

From (1) $\beta_0 = \bar{Y}_n - \beta_1 \bar{x}$ sub into (2)

get

$$\sum_{i=1}^{n} x_i Y_i = (\bar{Y}_n - \beta_1 \bar{x}) \overbrace{\sum_{i=1}^{n} x_i}^{n\bar{x}} + \beta_1 \sum_{i=1}^{n} x_i^2$$

$$= n \bar{x}_n \bar{Y}_n - \beta_1 n \bar{x}^2 + \beta_1 \sum_{i=1}^{n} x_i^2$$

$$\Rightarrow \sum_{i=1}^{n} x_i Y_i - n \bar{x}_n \bar{Y}_n = \beta_1 \left(\sum_{i=1}^{n} x_i^2 - n \bar{x}^2 \right)$$

$$\Rightarrow \beta_1 = \frac{\sum_{i=1}^{n} x_i Y_i - n \bar{x}_n \bar{Y}_n}{\sum_{i=1}^{n} x_i^2 - n \bar{x}^2} = \frac{\sum_{i=1}^{n} (x_i - \bar{x}_n)(Y_i - \bar{Y}_n)}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

2-variable 2nd derivative test...
wait

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}, \quad \hat{\beta}_0 = \bar{Y}_n - \hat{\beta}_1 \bar{x}$$

Unbiased? First

$$E(Y_i) = E(\beta_0 + \beta_1 x_i + \varepsilon_i) = \beta_0 + \beta_1 x_i + \underbrace{E(\varepsilon_i)}_{0}$$

$$E(\bar{Y}) = E\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(Y_i)$$

$$= \frac{1}{n}\sum_{i=1}^{n}(\beta_0 + \beta_1 x_i) = \frac{1}{n}\left(\sum_{i=1}^{n}\beta_0 + \beta_1\sum_{i=1}^{n} x_i\right)$$

$$= \frac{1}{n}\left(n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i\right) = \beta_0 + \beta_1 \bar{x}$$

$$E(\hat{\beta}_1) = E\left(\frac{\sum_{i=1}^{n}(x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)$$

$$= \frac{\sum_{i=1}^{n}(x_i - \bar{x})E(Y_i - \bar{Y})}{\sum(x_i - \bar{x})^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})\left(E(Y_i) - E(\bar{Y})\right)}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$= \frac{\sum_{i=1}^{n}(x_i - \bar{x})\left(\beta_0 + \beta_1 x_i - (\beta_0 + \beta_1 \bar{x})\right)}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

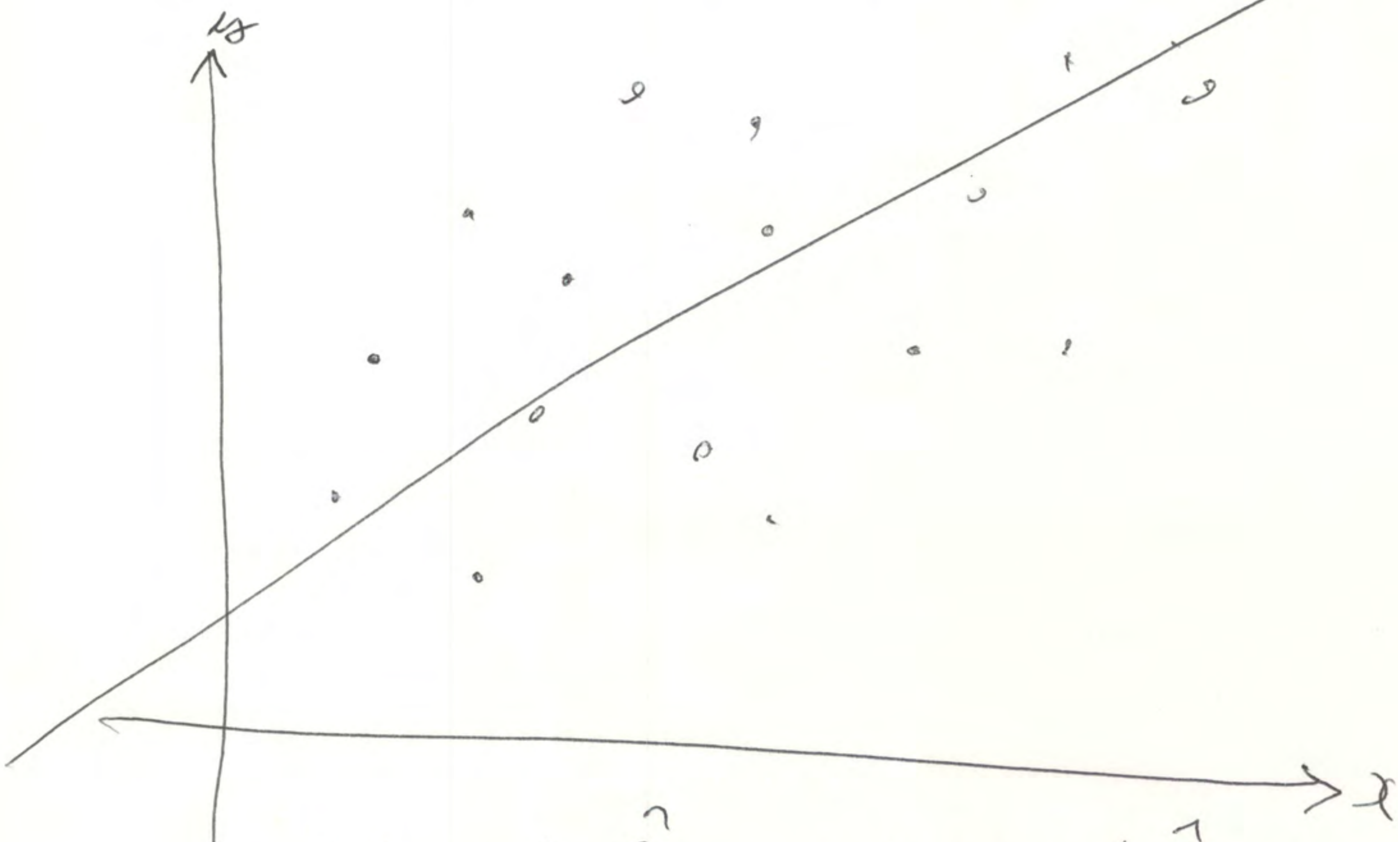
$$= \frac{\sum_{i=1}^{n}(x_i - \bar{x})\left(\beta_1(x_i - \bar{x})\right)}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \beta_1 \cancel{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

$= \beta_1$ unbiased

$$E(\hat\beta_0) = E(\bar Y - \hat\beta_1 \bar x)$$

$$= E(\bar Y) - \bar x E(\hat\beta_1)$$

$$= \beta_0 + \cancel{\beta_1 \bar x} - \cancel{\bar x \beta_1} = \beta_0 \quad \text{unbiased}$$



Slope is $\hat\beta_1$, intercept $\hat\beta_0$

They minimize $Q = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$

sum of squared vertical distances from point to line.