AMS131 25 Aug 17 ①

this time: important continuous distributions;

next time: large-sample theory

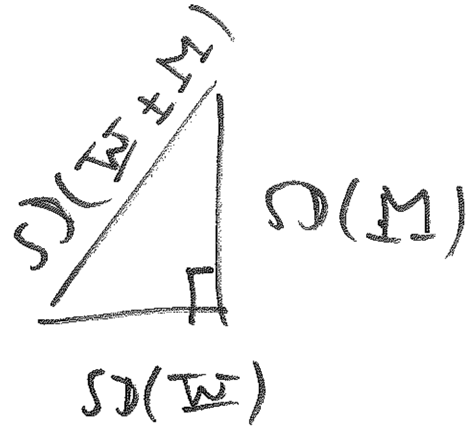
finish DS ch. 5, read DS ch. 6

canvas will shut 1 minute before midnight on Sun 27 Aug 2017 for take-home test 1

THT 2 canvas shut date will be 11.59 pm Wed 30 Aug 17

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

$$\sim N\left(\mu , \frac{\sigma^2}{n} \right)$$



$$E(\bar{X}_n) = E\left(\frac{1}{n} \sum_{i=1}^{n} X_i \right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i \right)$$

$$= \frac{1}{n} \sum_{i=1}^{n} E(X_i) \overset{(I)}{=} \frac{1}{n} \sum_{i=1}^{n} \mu = \mu$$

$$V(\bar{X}_n) = V\left(\frac{1}{n} \sum_{i=1}^{n} X_i \right) = \frac{1}{n^2} V\left(\sum_{i=1}^{n} X_i \right)$$

$$\frac{1}{n^2} V\left(\sum_{i=1}^{n} X_i\right) \overset{\text{I}}{=} \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) \overset{\text{ID}}{=} \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 \quad ②$$

$$\boxed{V(\bar{X}_n) = \frac{\sigma^2}{n}} \qquad \boxed{SD(\bar{X}_n) = \frac{\sigma}{\sqrt{n}}}$$

$\cancel{\theta = \bar{X}_n} \quad \cancel{E(\bar{X}_n) = \mu} \qquad \hat{\mu} = \bar{X}_n$

$E(\bar{X}_n) = \mu$

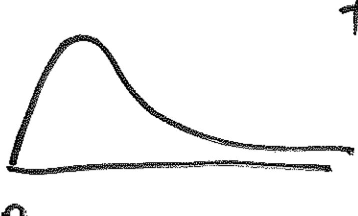
(unbiased)

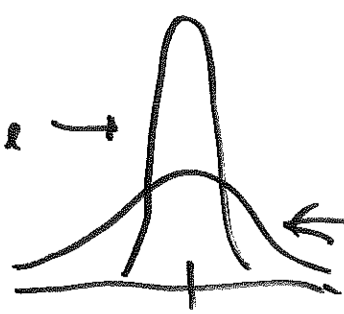
$E(\hat{\mu}) = \mu$

$SE(\hat{\mu}) = \frac{\sigma}{\sqrt{n}}$

uncertainty about μ

(9.45)



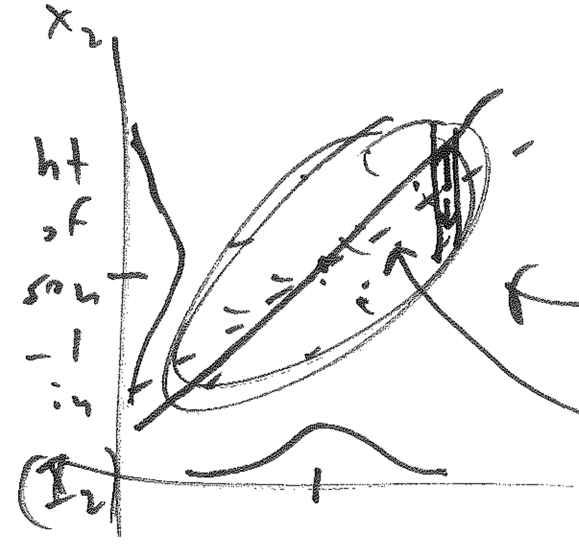


$X \sim N(\mu, \sigma^2)$

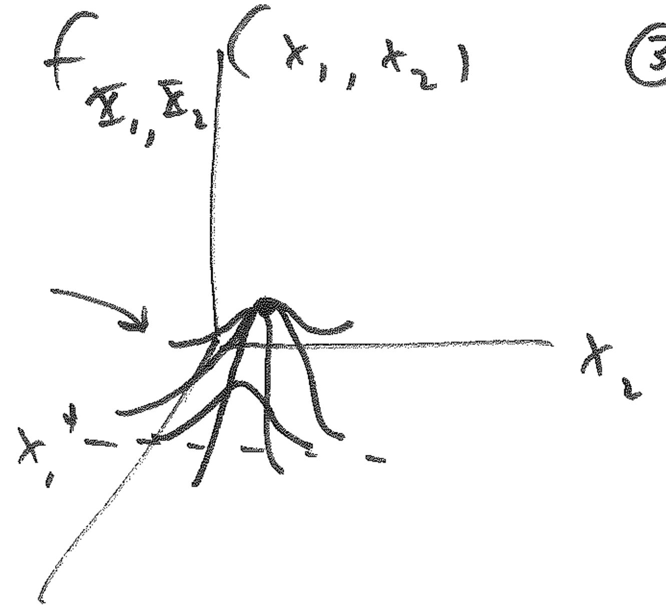
e^{X}

e^{Normal}

Galton



bivariate normal



ht. of father (X_1)

slope of regression line (for predicting X_2 from X_1).

$$
\begin{array}{c|c}
x_1 & x_2 \\
\end{array}
\quad
\begin{pmatrix}
x_{11} & x_{12} \\
\hline
x_{21} & x_{22} \\
\vdots & \vdots \\
x_{n1} & x_{n2}
\end{pmatrix}
\quad n
$$

mean $\bar{x}_1$ | $\bar{x}_2$

SD s_1 s_2

corr r

(prediction equation)

$$\hat{x}_2 = \hat{\beta}_0 + \hat{\beta}_1 x_1$$

$$\hat{\beta}_1 = r \frac{s_2}{s_1}$$

$$\hat{\beta}_0 = \bar{x}_2 - \hat{\beta}_1 \bar{x}_1$$

$$\hat{x}_2 = \bar{x}_2 - \hat{\beta}_1 \bar{x}_1 + \hat{\beta}_1 x_1 = \bar{x}_2 + \hat{\beta}_1 (x_1 - \bar{x}_1)$$

$$\hat{x}_2 = \bar{X}_2 + \hat{\beta}_1 (x_1 - \bar{x}_1)$$

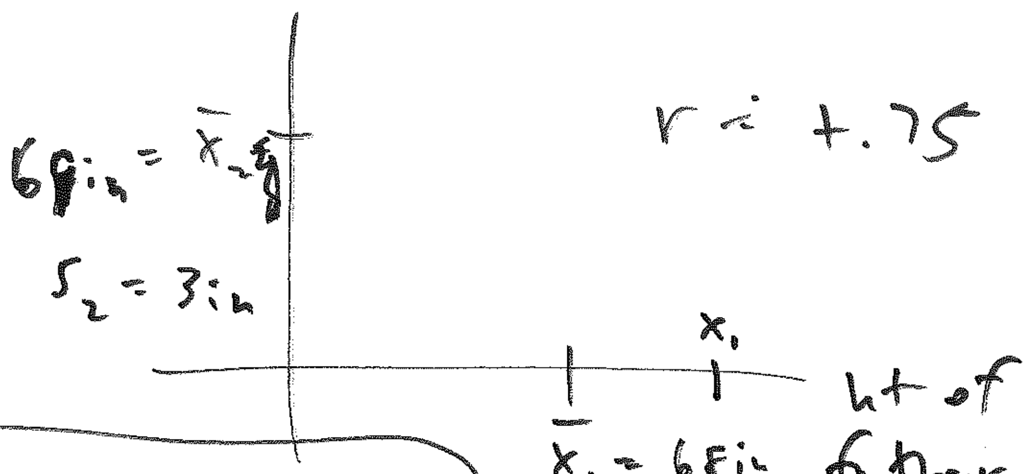
$$\bar{X}_2 \downarrow \mu_2 \qquad \hat{\beta}_1 \downarrow r\frac{s_2}{s_1} \downarrow \rho\frac{\sigma_2}{\sigma_1} \qquad \bar{x}_1 \downarrow \mu_1$$

$$E(X_2 \mid x_1) =$$

$\hat{x}_2$ trying to estimate $\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1)$

(10.57)

$r \doteq +.75$



$$\boxed{\hat{x}_2 = \bar{x}_2 + r\frac{s_2}{s_1}(x_1 - \bar{x}_1)}$$

$s_1 = 3$ in

typical

size of prediction error from regression line $= s_2\sqrt{1-r^2} \doteq 2.0$ in population

all possible measurements

X

N

mean μ

pop. density

sample

the observed measurements

rv

data values

X_1 x_1

$\vdots$ $\vdots$

n

at random

IID

X_n x_n

mean $\bar{X}_n$

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