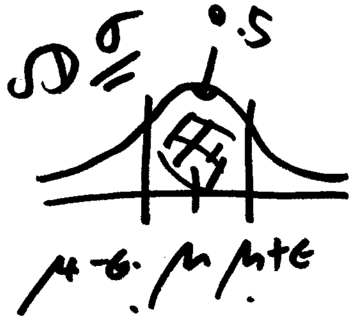


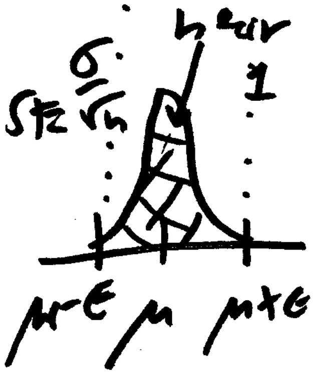
$(X_i \mid \mu, \sigma^2) \overset{IID}{\sim}$ ~~PMF~~ ~~PDF~~? (population) $N(\mu, \sigma^2)$

STAT 131
4 Jun 20

PDF of $\bar{X}_n = X_1$, $n = 1$

$$E(X_i) = \mu$$
$$V(X_i) = \sigma^2 < \infty$$
$$(i = 1, \ldots, n)$$

DD office ①
1.5-hour session

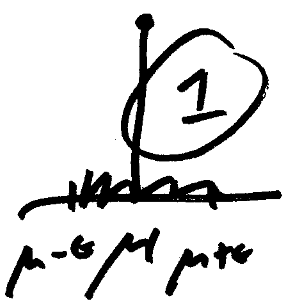


PDF of $\bar{X}_n$, $n = 10$

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

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

$$V(\bar{X}_n) =$$



PDF of $\bar{X}_n$, $n \to \infty$

$$\boxed{SE(\bar{X}_n) = SD(\bar{X}_n) = \frac{\sigma}{\sqrt{n}}}$$

$$\lim_{\cancel{n \to \infty}} P(|\bar{X}_n - \mu| \le \epsilon) = 1$$

fix small $\epsilon > 0$

WLLN

A sequence of r.v. $W_1, W_2, \ldots$ converges in probability to a constant c iff for all $\epsilon > 0$ $\lim_{n \to \infty} P(|W_n - c| \le \epsilon) = 1$

$$y_i = \underline{\underline{\theta + b}} + e_i \quad (i = 1, \dots, n)$$

θ: truth; b: bias; e_i: r.v.

wLOG

$(e_i \mid \sigma^2) \stackrel{IID}{\sim} N(\underline{0}, \sigma^2)$

assume $0 < \sigma < \infty$

$E(y_i) =$

HHT's $I(a)$

$$(y_i \mid \theta, b, \sigma^2) \stackrel{IID}{\sim} N(\theta + b, \sigma^2) \quad (i = 1, \dots, n)$$

$$\bar{y}_n = \frac{1}{n} \sum_{i=1}^{n} y_i$$

WLLN: $\bar{y}_n \xrightarrow{P} E(y_i)$

$$E(y_i) = E(\theta + b + e_i) = \theta + b + \cancel{E(e_i)}^{0}$$

$$= \theta + b$$

WLLN: $\bar{y}_n \xrightarrow{P} \theta + b$

and therefore $\boxed{\bar{y}_n \xrightarrow{P} \theta \text{ iff } b = 0}$

i.e., measuring process is unbiased

$$RMSE(\bar{X}_n) = \sqrt{E(\bar{X}_n - \theta)^2}$$

$$MSE(\bar{X}_n) = E\left[(\bar{X}_n - \theta)^2\right]$$

$$= E\left(\bar{X}_n^2 - 2\theta\bar{X}_n + \theta^2\right)$$

$$= E(\bar{X}_n^2) + E(-2\theta \cdot \bar{X}_n) + E(\theta^2)$$

$$= E(\bar{X}_n^2) - 2\theta E(\bar{X}_n) + \theta^2$$

$$= E(\bar{X}_n^2) - 2\theta(\theta + b) + \theta^2$$

$$\boxed{E(\bar{X}_n) = \theta + b}$$

$$V(\bar{X}_n) = E(\bar{X}_n^2) - \left(E(\bar{X}_n)\right)^2$$

$$\text{So } E(\bar{X}_n^2) = V(\bar{X}_n) + \left[E(\bar{X}_n)\right]^2$$

$$= \frac{\sigma^2}{n} + (\theta + b)^2$$

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

$$MSE(\bar{\Sigma}_n) = \frac{\sigma^2}{n} + (\theta + b)^2$$

$$-2\theta^2 - 2\theta b + \theta^2$$

$$= \frac{\sigma^2}{n} + \cancel{\theta^2} \cancel{+ 2\theta b} + b^2 - \cancel{2\theta^2}$$

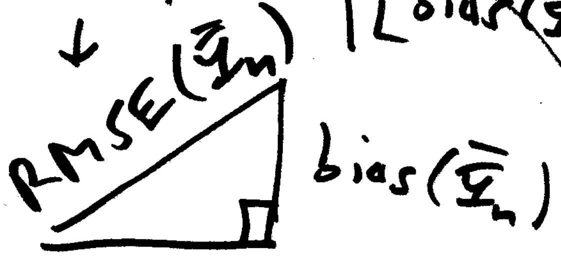
$$\cancel{-2\theta b} \cancel{+ \theta^2}$$

$$MSE(\bar{\Sigma}_n) = \frac{\sigma^2}{n} + b^2$$

$$RMSE(\bar{\Sigma}_n) = \sqrt{\frac{\sigma^2}{n} + b^2} \quad ✓$$

$$= \sqrt{[SE(\bar{\Sigma}_n)]^2 + [bias(\bar{\Sigma}_n)]^2} = \sqrt{\left(\frac{\sigma}{\sqrt{n}}\right)^2 + b^2}$$

want small ↓



$SE(\bar{\Sigma}_n) = \frac{\sigma}{\sqrt{n}}$

↑ sampling uncertainty

to make RMSE($\bar{\Sigma}_n$) small, want n to be large, want b to be close to 0

$$\bar{e}_n = \frac{e_1 + e_2 + \dots + e_n}{n}$$

weight (oz). ⑤

$$\begin{bmatrix} 15.97 \\ 16.01 \\ \vdots \\ 16.01 \end{bmatrix} \uparrow n \downarrow$$

suppose butter-weighing is unbiased & truth = 16 oz = θ

		truth	+ bias +	random error
15.97	=	16.0	+ 0	+ (−0.03)
16.01	=	16.0	+ 0	+ (+0.01)
⋮		⋮	⋮	⋮
16.01	=	16.0	+ 0	+ (+0.01)

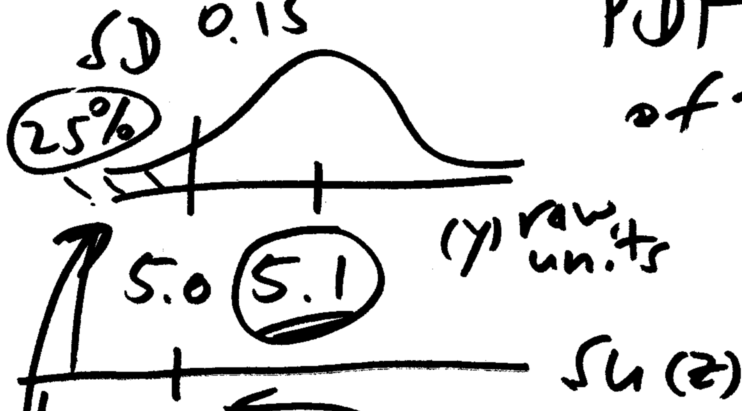
$$\bar{y}_n = \theta + 0 + \frac{(-0.03) + (+0.01) + \dots + (+.01)}{n}$$

$\boxed{E(e_i) = 0}$

$SD(e_i) = \sigma = 0.15$

$b = 0$

$y_i = 5.1 + 0 + e_i$ ← IID $N(0, \sigma^2)$



PDF of $Y_1 = 5.1 + 0 + e_1$

$e_1 \sim N(0, \sigma^2)$

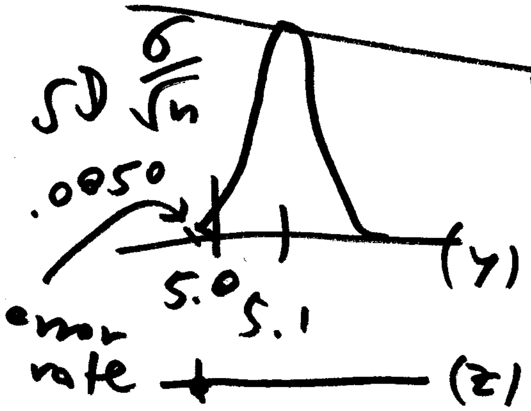
$\sigma = 0.15$

$$\frac{y-\mu}{\sigma} = \frac{5.0-5.1}{0.15} = \frac{-0.1}{0.15} = -\frac{2}{3} = -0.67$$

$\Phi(-0.67)$

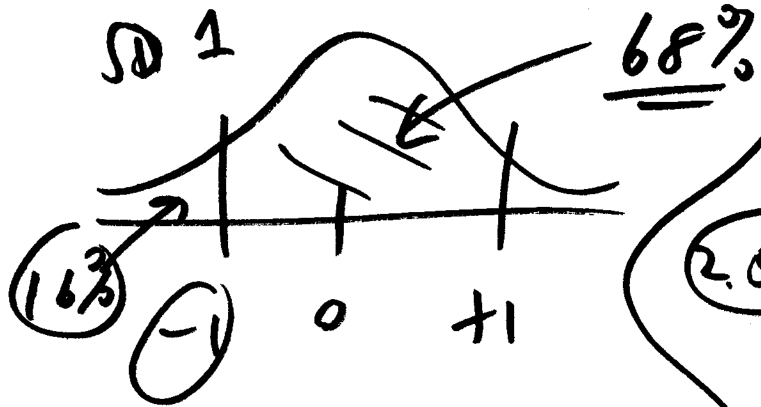
P(we say threatened based on $\bar{Y}_n$, n=1 | truth = 5.0 (not threatened))

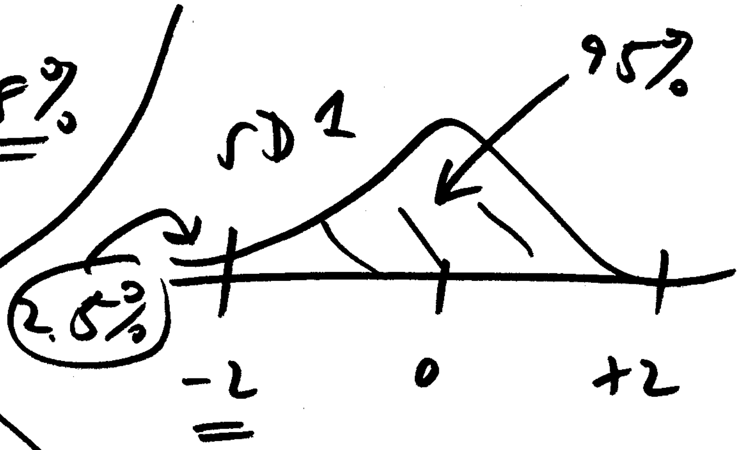
≐ 25% ← misclassification (error) rate

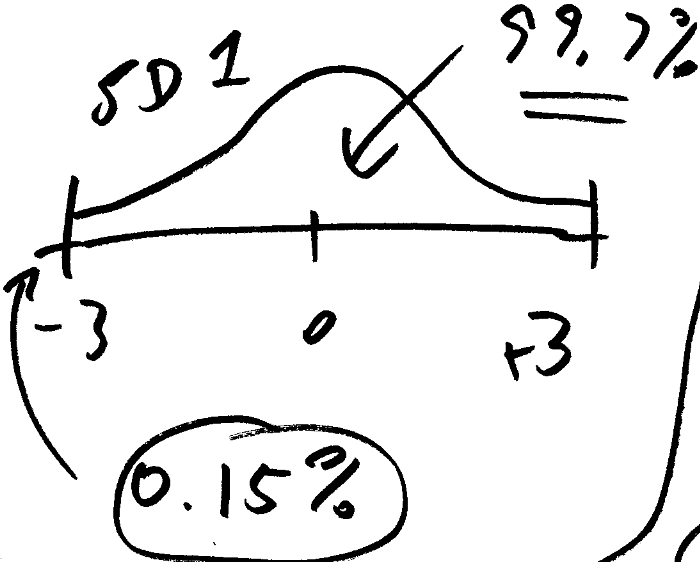


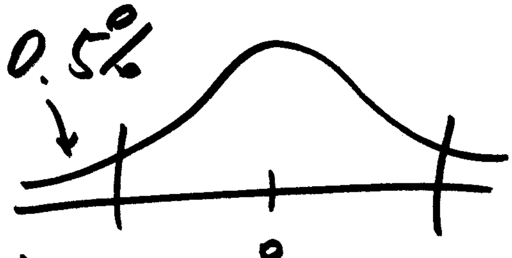
PDF of $\bar{Y}_n$, $n > 1$

$$\frac{y-\mu}{\sigma/\sqrt{n}} = \frac{5.0-5.1}{0.15/\sqrt{n}} = \Phi^{-1}(.005)$$





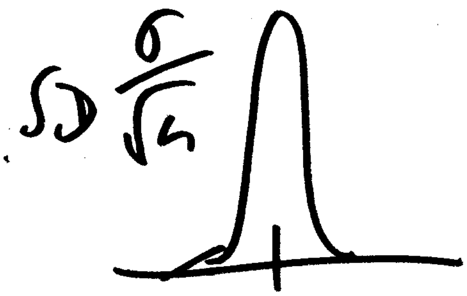




between

-2 and -3

-2.575



PDF of $\bar{X}_n$

5.0

$$-2.575 = \frac{5.0 - 5.1}{0.15/\sqrt{n}}$$