# CS 70 Discussion 7B

CLT:

Let $X_1, X_2 ... X_n$ be a sequence of iid RVs with any distribution. Let each $X_i$ have common mean $\mu$ and variance $\sigma^2$. We can define:

$$Z_n = \frac{X_1 + ... + X_n - n\mu}{\sigma\sqrt{n}}$$

$\mu = \frac{1}{2}$

such that $Z_n$ has the same distribution as the standard normal RV (0,1).

1 (a) $\underbrace{P(|X-\mu| \geq \varepsilon)} \leq \frac{Var(X)}{\varepsilon^2} = \boxed{\frac{\sigma^2}{\varepsilon^2}}$

(b) $P(|X-\mu| < \varepsilon)$ is the same as $P\{\mu \in (X-\varepsilon, X+\varepsilon)\}$

$|X-\mu| < \varepsilon$

$X-\mu < \varepsilon$ and $X-\mu > -\varepsilon$

$X < \mu+\varepsilon$ and $X > \mu-\varepsilon \Rightarrow \boxed{\mu-\varepsilon < X < \mu+\varepsilon}$

(c) $\varepsilon = ?$ $\quad P(\mu \in (X-\varepsilon, X+\varepsilon)) > 95\%$

$P(|X-\mu| > \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} \leq 5\%$

$\mu \in (X - \sigma\sqrt{20},$

$$20\left(\frac{\sigma^2}{\varepsilon^2}\right) = (0.05)^{20}$$

$$\varepsilon^2 = 20\sigma^2$$

$$\boxed{\varepsilon = \sigma\sqrt{20}}$$

$\boxed{X \pm \sigma\sqrt{20}) \text{ w/ } 95\% \text{ confidence}}$

[2] Use Chebyshev's

$1-\delta$ confidence interval $\Rightarrow$ $\delta$ prob. we are outside interval

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

actually an <u>estimator</u> for $\lambda$

$$\mathbb{E}[\bar{X}] = \mathbb{E}\left[\frac{\sum_{i=1}^{n} X_i}{n}\right] = \frac{1}{n}\underline{\sum_{i=1}^{n}\mathbb{E}[X_i]}$$

$$= \frac{\cancel{n}\lambda}{\cancel{n}} = \underline{\lambda}$$

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

$$= \frac{1}{n^2}\underline{\sum_{i=1}^{n}}\operatorname{Var}(X_i) = \frac{\cancel{n}\lambda}{n^{\cancel{2}}} = \frac{\lambda}{n}$$

b/c each of the $X_i$ are iid

lim $n \to \infty$

$$P(|\bar{X} - \lambda| \geq \varepsilon) \leq \frac{\operatorname{Var}(\bar{X})}{\varepsilon^2} = \frac{\lambda}{n\varepsilon^2} \leq \frac{2}{n\varepsilon^2}$$

$\underline{\lambda \leq 2}$

<u>Find a $1-\delta$ conf. interval</u> $\Rightarrow \leq \delta$ prob. of being <u>outside</u> that conf. interval

$$\frac{2}{n\varepsilon^2} \leq \delta$$

$$\frac{2}{\delta n} \leq \frac{\delta n \varepsilon^2}{\delta n}$$

$$\varepsilon^2 \geq \frac{2}{\delta n}$$

$$\boxed{\varepsilon \geq \sqrt{\frac{2}{\delta n}}}$$

Idea: if we want to be '$1-\delta$' % confident that our conf. interval captures $\lambda$, then we should set the width of our confidence interval to be $\varepsilon \geq \sqrt{\frac{2}{\delta n}}$

$$\text{Var}(\bar{X}) = \text{Var}\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n^2}\text{Var}\Big(\sum_{i=1}^{n} \underbrace{X_i}_{\text{iid}}\Big) = \frac{1}{n^2}\sum_{i=1}^{n}\text{Var}(X_i)$$

$$= \frac{1}{n^2} n(\lambda) = \frac{\lambda}{n}$$

$\boxed{3}$ CLT says that for $Z_n = \dfrac{X_1 + \dots + X_n - n\mu}{\sigma\sqrt{n}}$, that $Z_n$ converges to the standard normal distribution as $n \to \infty$.

In particular, $\lim_{n\to\infty} P(Z_n \leq z) = \Phi(z)$ where $\Phi$ is the CDF of the standard normal RV. $\mu = \mathbb{E}[X_i]$

Define $X_i$ to be the outcome of the $i^{th}$ coin toss

$$X_i = \begin{cases} 1 & \text{if heads} \\ 0 & \text{if tails} \end{cases}$$

Bernoulli RV

$$\mathbb{E}[X_i] = \frac{1}{2}$$

$$\text{Var}(X_i) = p(1-p) = \frac{1}{2}\left(\frac{1}{2}\right) = \underbrace{\frac{1}{4}}_{\sigma^2}$$

$$\frac{\overbrace{X_1 + \dots + X_n}^{} - n\underbrace{\mu}_{\frac{1}{2}}}{\sigma\sqrt{n}}$$ is distributed as $N(0,1)$

Let $Y = X_1 + \dots + X_{100}$ (so $Y$ is the number of RV)

$$\boxed{\frac{Y - n\mu}{\sigma\sqrt{n}}} \sim N(0,1)$$

$n = 100$

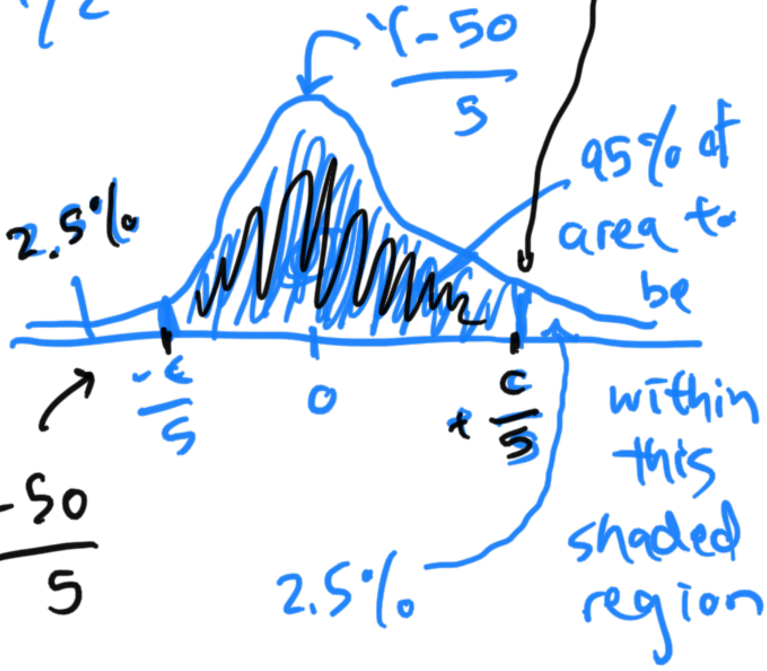
$\mu = 0.5$

$\sigma = 1/2$

$\Phi(x) = .975$

$x = \Phi^{-1}(.975)$

$$\frac{Y-50}{5} \sim N(0,1)$$




$$P(|Y-50| \le c) \ge 0.95$$

$$P\left(\frac{|Y-50|}{5} \le \frac{c}{5}\right) \ge 0.95$$

$$+\frac{c}{5} = \Phi^{-1}(.975)$$

$50 \qquad 50 \pm 9.8$

$$\frac{c}{5} = 1.96$$

$$\boxed{c = 9.8}$$

$\ge 59.8$ (60)

$\le 40$

If we flip coin 100 times and observe $\ge 60$ heads or $\le 40$ heads, we can declare coin to be biased and be correct 95% of the time.