FoDA: L7

Concentration of Measure

CLT

$E[X_i] = \mu$

$Var[X_i] = \sigma^2$

n R.V. $X_1, X_2 \ldots X_n \underset{iid}{\sim} f$

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

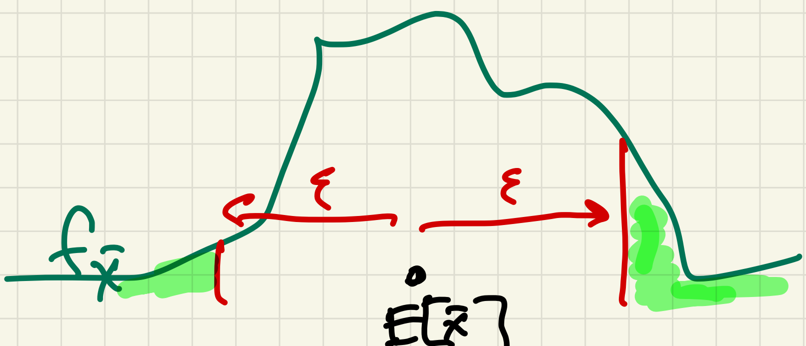
- $E[\bar{X}]$
- $Var[\bar{X}] = \sigma^2/n$
- converges $N(\mu, \sigma^2/n)$

Probably Approximatly Correct (PAC)

$$\Pr\left[\,|\bar{x} - E[\bar{x}]| \geq \varepsilon\,\right] \leq \delta$$

error tolerance (ε)

probability of failure (δ)



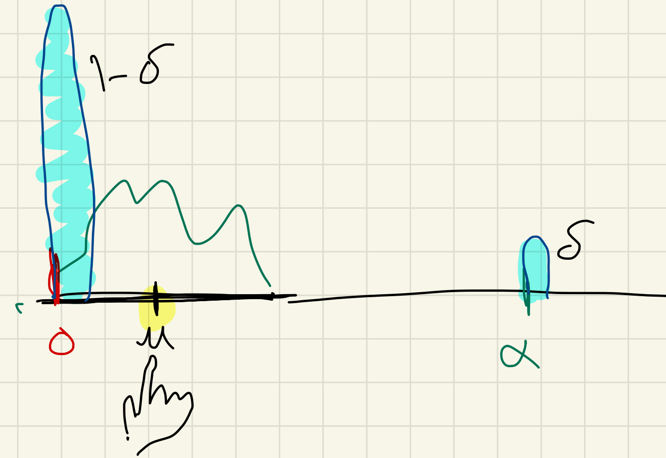
Markov Inequality

R.V. X

- $E[X]$
- $X \geq 0$

for any $\alpha > 0$

$$\boxed{\Pr[X > \alpha] \leq \frac{E[X]}{\alpha}} = \delta$$



$$E[X] = \cancel{(1-\delta)} \cdot 0 + \delta \cdot \alpha$$

$$\delta = \frac{E[X]}{\alpha}$$

$$\varepsilon = \alpha - E[X] \qquad \delta = \frac{E[X]}{\alpha}$$

$$\Pr[X - E[X] > \varepsilon] \leq \frac{E[X]}{(\varepsilon + E[X])} = \delta$$

R.V. R rain in SLC in June in mm

$R \geq 0$

$E[R] = 20$ mm

$$\Pr[R \geq 50\,\text{mm}] \leq \frac{E[R]}{50} = \frac{20}{50} = 0{,}4$$

Chebyshev Inequality

R.V. X

- $E[X] = \mu$
- $Var[X] = \sigma^2$ for any $\varepsilon > 0$

$$\boxed{\Pr\left[\,|X - E[X]| \geq \varepsilon\,\right] \leq \frac{Var[X]}{\varepsilon^2}} = \delta$$

R SLC Rain June

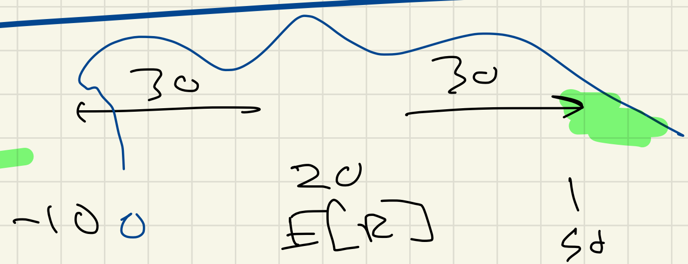
$E[R] = 20\,mm$

$Var[R] = 9\,mm^2$

$\Pr[R \geq 50]$

$\leq \Pr\left[\,|R - E[R]| \geq 30\,\right]$

$\leq \frac{9\,mm^2}{(30\,mm)^2} = 0.01$



Chebyshev

$X_1, X_2, \ldots, X_n$ $\qquad X_i \sim f$

$E[X_i] = \mu$ $\qquad E[\bar{X}] = \mu$

$Var[X_i] = \sigma^2$ $\qquad Var[\bar{X}] = \sigma^2 / n$

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

$$Pr\left[\left| \bar{X} - E[\bar{X}] \right| > \varepsilon \right] \leq \frac{Var[\bar{X}]}{\varepsilon^2} = \frac{\sigma^2}{n \varepsilon^2}$$

Fair die T R.V. roll 3

$$T_i = \begin{cases} 1 & \text{if } 3 \\ 0 & \text{o.w.} \end{cases}$$ w.p. 1/6

$T = \sum_i T_i$

Roll die $n = 120$ times

T = # times roll 3

$E[T] = 20$

$Var(T_i) = 5/36$

$$\Pr[T > 40] \leq$$

$$\Pr[|T - \underset{\substack{\shortparallel \\ E[T]}}{20}| \geq 20] \leq \frac{Var[T_i]}{n \cdot 20^2} = \frac{5/36}{120 \cdot 400} \leq 0.042$$

Chernoff-Hoeffding Inequality

$X_1, X_2, \ldots X_n$ RV. $\{X_i\} \underset{iid}{\sim} f$

$E[X_i] = \mu$

$X_i \in [a, b]$ $\Delta = b - a$

$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ for any $\varepsilon > 0$

$$\boxed{\Pr\left[\left|\bar{X} - E[\bar{X}]\right| \geq \varepsilon\right] \leq \overbrace{2\exp\left(-\frac{2\varepsilon^2 n}{\Delta^2}\right)}^{\delta}}$$

$\Delta = 1$ $\quad T_i = \begin{cases} 1 & \text{w.p. } 1/6 \\ 0 & \text{o.w} \end{cases}$ $\quad E[T_i] = 1/6 \quad \text{Var}[T_i] = 5/36$

$\bar{T} = \frac{1}{n}\sum_{i=1}^{n} T_i \qquad n = 120 \qquad T = \#\, 3s$

$= \bar{T} \cdot n$

$\Pr[n \cdot \bar{T} > 40] = \Pr[\bar{T} > 1/3] \leq 0.0026$

$= \Pr[\bar{T} - E[\bar{T}] > 1/6] \leq \Pr[|\bar{T} - E[\bar{T}]| \geq 1/6]$

C-H

$\leq \boxed{2 \cdot \exp\left(\frac{-2\varepsilon^2 \cdot n}{\Delta^2}\right)} = 2 \cdot \exp\left(-\frac{2(1/6)^2 \cdot 120}{1}\right)$

$= 2\exp\left(-\frac{1}{18} \cdot 120\right) = 2 \cdot \exp\left(-\frac{20}{3}\right) \approx 0.0026$

$$2 \cdot \exp\left(-\frac{2\varepsilon^2 n}{\Delta^2}\right) = \delta$$

C-H

Given error tolerance ε (= 0.1)

prob. of failure δ (= 0.05)

$$\exp\left(-\frac{2\varepsilon^2 n}{\Delta^2}\right) = \delta/2$$

$$\exp\left(\frac{2\varepsilon^2 n}{\Delta^2}\right) = 2/\delta$$

$$\frac{2\varepsilon^2 n}{\Delta^2} = \ln(2/\delta)$$

$$n = \frac{\Delta^2}{2\varepsilon^2} \ln(2/\delta)$$

Cheb

$$\frac{\sigma^2}{n\varepsilon^2} = \delta$$

$$\delta n = \frac{\sigma^2}{\varepsilon^2}$$

$$n = \frac{\sigma^2}{\varepsilon^2 \delta}$$

Markov	1 R.V	$E[X]$ $X \geq 0$
Cheb	1 or R.V. n	$E[X_i]$ $Var[X_i]$
C-H	n R.V.	$E[X_i]$ Δ st $X_i \in [a,b]$ $\Delta = b-a$