FoDA L7 : Concentration of Measure

Probably Approx Correct (PAC)

unknown function (pdf) f

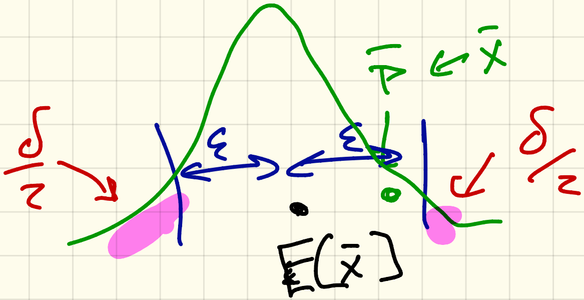
$\{X_1, \ldots X_n\}_{iid} f$

$$\bar{X} = \frac{1}{n} \sum_i X_i$$

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

error tolerance

probability of failure



Markov Inequality

R.V. X

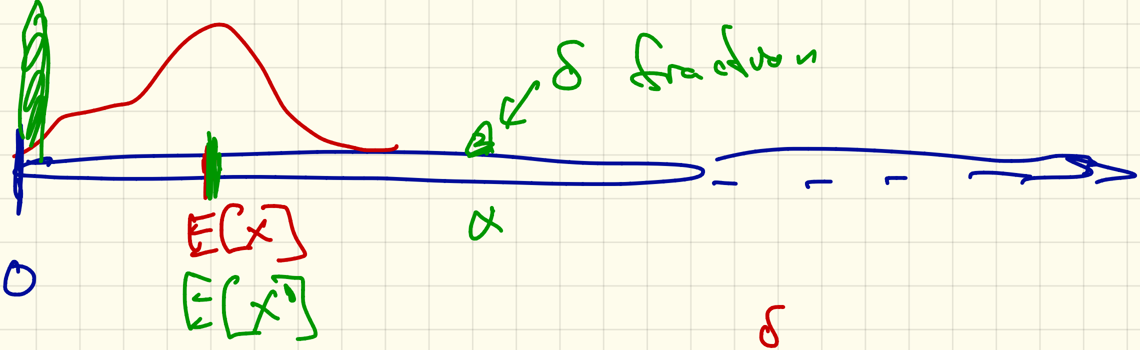
(a) $X \geq 0$ (b) $E[X]$ known

any $\alpha > 0$

$$\boxed{P_r[X > \alpha] \leq \frac{E[X]}{\alpha}}$$

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

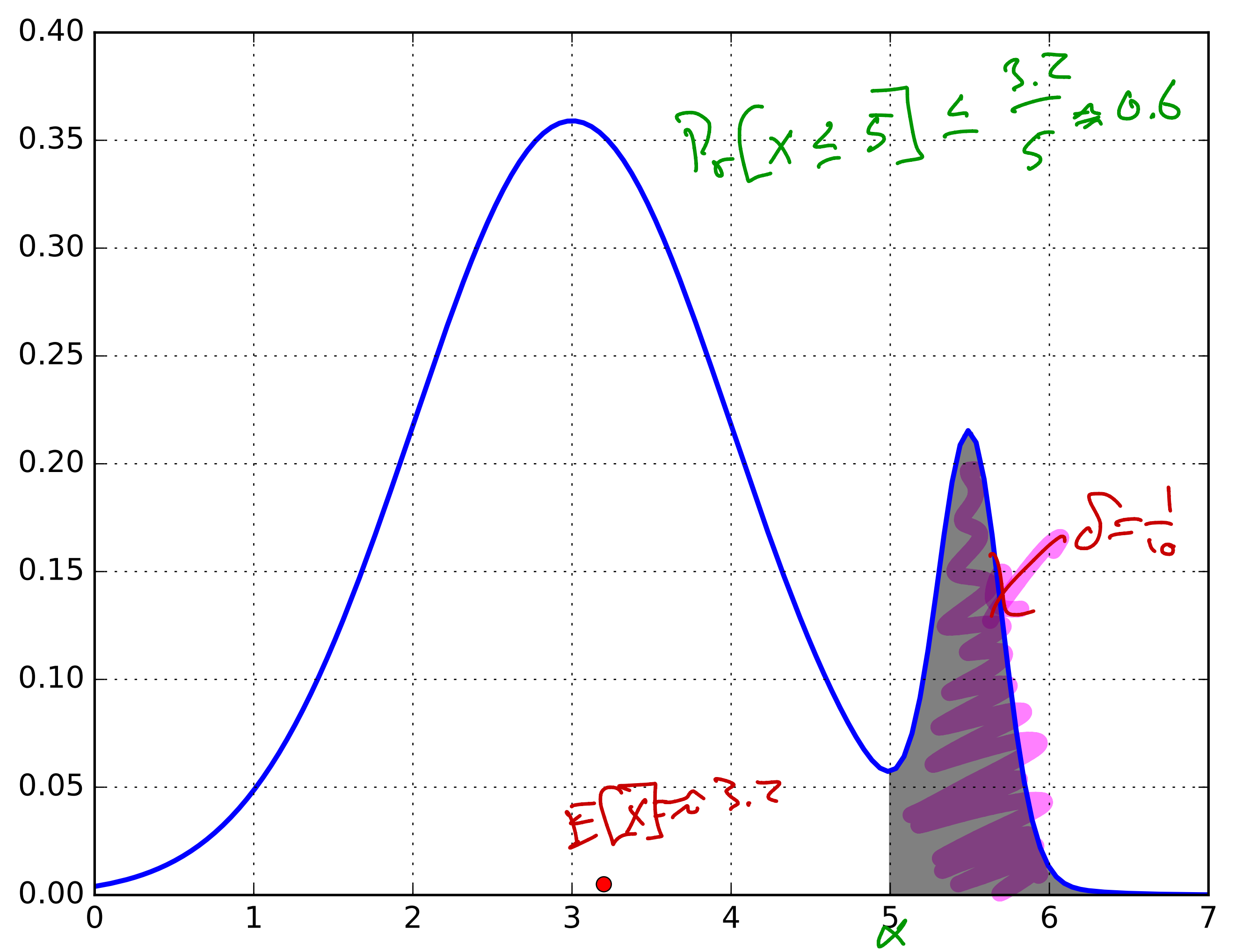
$$P_r[X - E[X] > \varepsilon] \leq \delta = \frac{E[X]}{\varepsilon + E[X]}$$

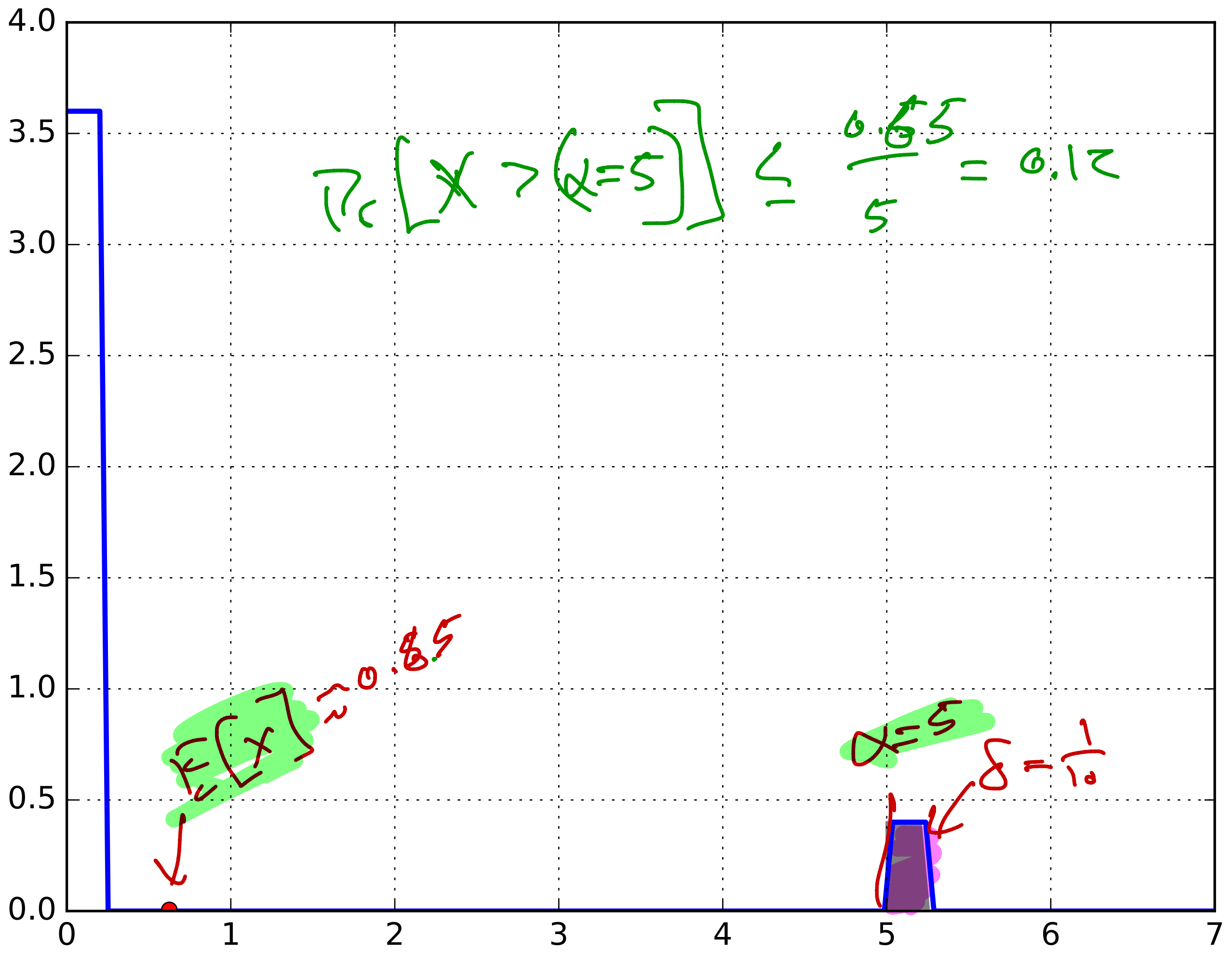


$$E[X] = 0 \cdot \underbrace{\Pr[X=0]}_{1-\delta} + \alpha \cdot \overbrace{\Pr[X=\alpha]}^{\delta}$$

$$= 0 \cdot (1-\delta) + \alpha \cdot \delta = \alpha \cdot \delta$$

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


0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
1
2
3
4
5
6
7
α
Pr[X ≤ 5] ≤ 3.2/5 ≈ 0.6
E[X] = 3.2
δ = 1/10



4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
1
2
3
4
5
6
7
$\Pr\left[X > (\alpha = 5)\right] \le \frac{0.65}{5} = 0.13$
$E[X] \approx 0.65$
$\alpha = 5$
$\delta = \frac{1}{10}$

June expect rain of SLC

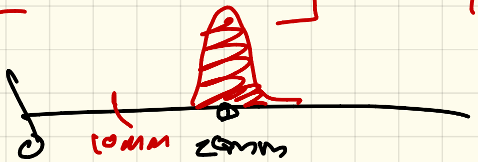
R.V. R

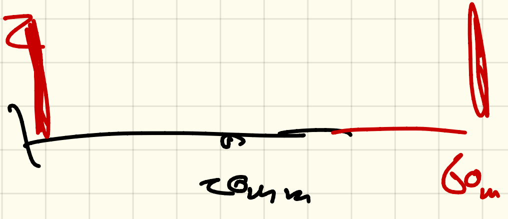
$$E[R] = 20 \text{ mm}$$

Prob Rain in June in SLC $\geq$ 50mm

$$\Pr[R > 50\text{m}] \leq \frac{E[R]}{50 \text{ mm}} = \frac{20}{50} = 0.4$$

$$\Pr[R > 10 \text{ mm}] \leq \frac{20\text{m}}{10\text{m}} = 2$$





Chebyshev Inequality

R.V. X

(a) $E[X]$ (b) $Var[X]$

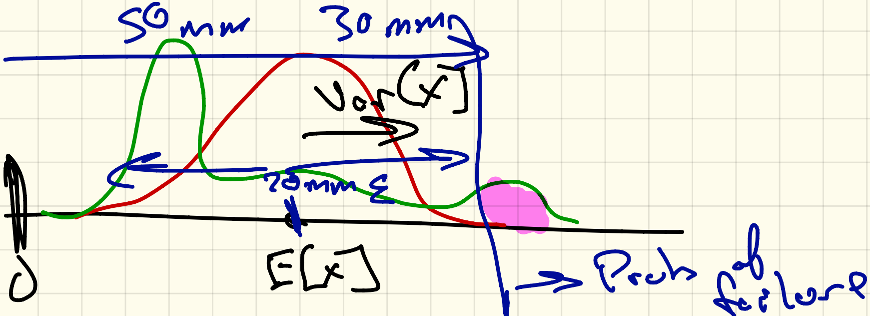
for any $\varepsilon > 0$

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

error

PAC

$$\delta = \frac{Var[X]}{\varepsilon^2}$$



Rainfall June in SLC

R.V. R

$E[R] = 20\,mm$ $\qquad$ $Var[R] = 9\,mm^2$

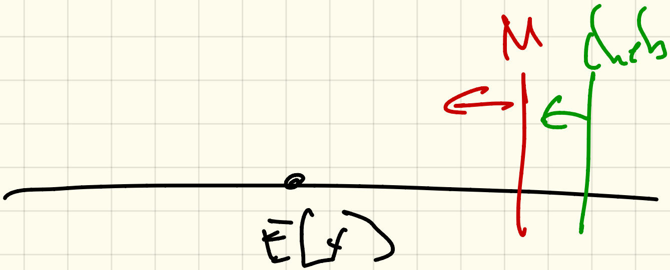
$90\,mm^2$

$$Pr[R \geq 50\,mm] = Pr[R - \overset{E[R]}{20\,mm} \geq 30\,mm]$$

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

$$\leq \frac{Var[R]}{30^2} = \frac{9}{900} = 0.01$$

$$\frac{90}{900} = 0.10$$



RVs $X_1, X_2, \ldots X_n \overset{\wedge}{\text{iid}}\, f$

$$\bar{X} = \frac{1}{n}\sum_i X_i \qquad \mathrm{Var}[\bar{X}] = \frac{\mathrm{Var}[X_i]}{n}$$

$$\Pr\left[\,|\bar{X} - E[X_i]| \geq \varepsilon\,\right] \leq \frac{\mathrm{Var}[\bar{X}]}{\varepsilon^2}$$

$$\leq \frac{\mathrm{Var}[X_i]}{n \cdot \varepsilon^2} = \delta$$

Given:
- $\mathrm{Var}[X_i]$
- error tolerance ε
- prob. failure δ

How many samples (n) are needed

$$n = \frac{\mathrm{Var}[X_i]}{\varepsilon^2 \delta}$$

Chernoff-Hoeffding Inequality

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

(a) $E[X_i]$ (b) $X_i \in [b, t]$ $\Delta = t - b$

Examp Rainfall $\in [10in, 60in]$ b t $\Delta = 60 - 10 = 50$

$\bar{X} = \frac{1}{n} \sum_i X_i$

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

δ

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

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

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

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

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

Given: $\varepsilon, \delta, \Delta$ How many samples (n) are needed

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

Roll Dice $n = 120$ times

R.V. $T = \#$ roll a 3

$$E[T] = 120 \cdot \frac{1}{6} = 20$$

$$\Pr[T \geq 40] \leq ?$$

R.V. X_i: $t = 1$ 1 if 3, $b = 0$ 0 if $\{1,2,4,5,6\}$

$$\bar{X} = \frac{1}{n}\sum_i X_i$$

$$T = \bar{X} \cdot 120$$

$$= \Pr\left[\bar{X} \geq \frac{1}{3}\right] \leq \Pr\left[\left|\bar{X} - E[\bar{X}]\right| \geq \frac{1}{6}\right]$$

$$\varepsilon = \frac{1}{3} - E(\bar{X}) = \frac{1}{3} - \frac{1}{6}$$

C-H

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

ε^2, $= n$

$$= 2\exp\left(\frac{-20}{3}\right) \leq 0.0026$$

Cheb $\mathrm{Var}[X_i] = \frac{5}{36}$

$$\mathrm{Var}[\bar{X}] = \frac{5}{36} \cdot \frac{1}{120}$$

$$\leq \frac{\mathrm{Var}[\bar{X}]}{(1/6)^2} = \frac{\frac{5}{36} \cdot \frac{1}{120}}{1/36} = \frac{5}{120} \approx 0.042$$