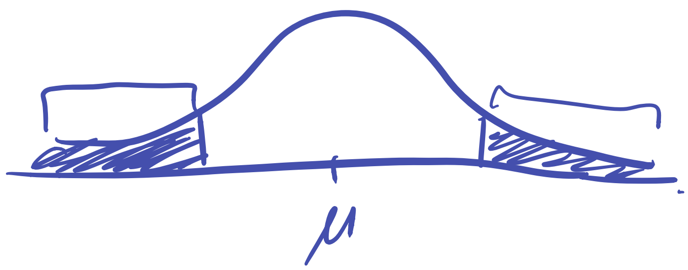
Dis 12A: Tail Bounds

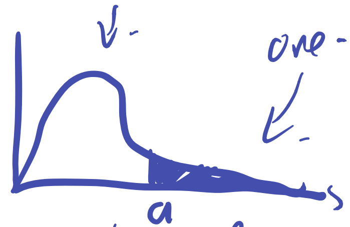
Riddle: 5 employees want to find out the sum of their salaries without ever giving away their own. How can they accomplish this?

chebyshev's

$$\boxed{\mathbb{P}(|X-\mu| > k)}$$



Markov's, one-sided



Markov's Inequality

Let X be a nonneg. r.v. w/ finite mean. Then for any $a > 0$,

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}.$$

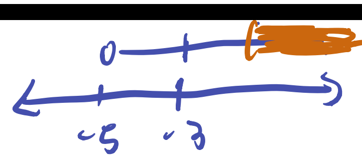
Chebyshev's Inequality (symmetric)

Let X be a r.v. w/ finite mean & variance. Then for any $k > 0$,

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq k) \leq \frac{Var(X)}{k^2}.$$

e.x. of symmetric distr. $\quad \mathbb{P}[X=x] = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

1 Inequality Practice



(a) X is a random variable such that $X > -5$ and $\mathbb{E}[X] = -3$. Find an upper bound for the probability of X being greater than or equal to -1.

$X > -5 \Rightarrow X' > 0$.

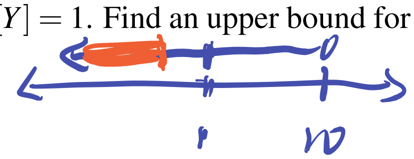
$X + 5 > 0 \Rightarrow$

Set $X' = X + 5$.

$\mathbb{P}[X \geq -1] = \mathbb{P}[X+5 \geq 4]$

$= \mathbb{P}[X' \geq 4] \leq \frac{\mathbb{E}[X']}{4} = \frac{-3+5}{4} = \frac{1}{2}$.

(b) Y is a random variable such that $Y < 10$ and $\mathbb{E}[Y] = 1$. Find an upper bound for the probability of Y being less than or equal to -1.



$Y < 10 \Rightarrow Y' > 0$

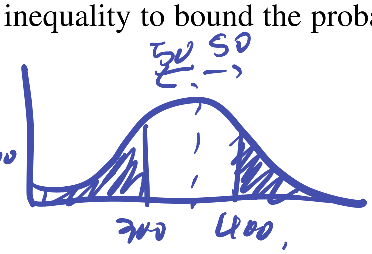
$10 - Y > 0 \Rightarrow$

$Y' = 10 - Y$.

$\mathbb{P}[Y \leq -1] = \mathbb{P}[0 \leq -1 - Y] = \mathbb{P}[11 \leq 10 - Y]$.

$= \mathbb{P}[Y' \geq 11] \leq \mathbb{E}[Y']/11 = \boxed{9/11}$

(c) You roll a die 100 times. Let Z be the sum of the numbers that appear on the die throughout the 100 rolls. Compute $\operatorname{Var}(Z)$. Then use Chebyshev's inequality to bound the probability of the sum Z being greater than 400 or less than 300.



Z = sum of the rolls $= Z_1 + \cdots + Z_{100}$

$\mathbb{E}[Z] = 100 \cdot \mathbb{E}[\text{one roll}] = 350$.

$\operatorname{Var}(Z_i) = \mathbb{E}[Z_i^2] - \mathbb{E}[Z_i]^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$.

$\mathbb{E}[Z_i^2] = \sum_{z=1}^{6} z^2\, \mathbb{P}(Z_i = z) = \frac{1}{6} \cdot \sum_{z=1}^{6} z^2 = \frac{91}{6}$.

$\operatorname{Var}(Z) = \operatorname{Var}(Z_1) + \operatorname{Var}(Z_2) + \cdots + \operatorname{Var}(Z_{100}) = 100 \cdot \left(\frac{35}{12}\right)$

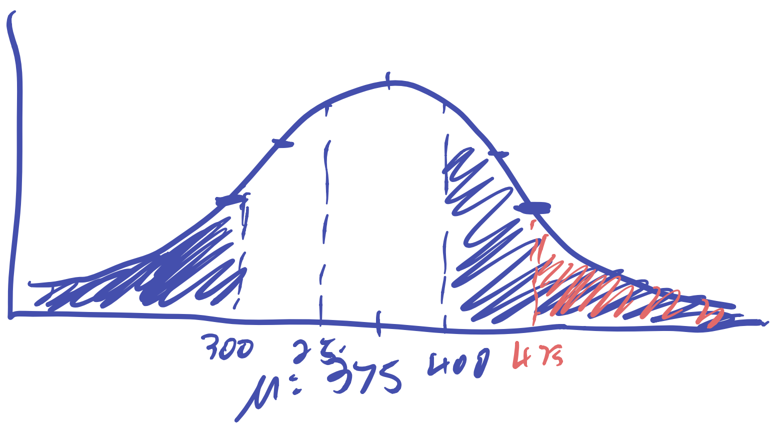
$\underline{Z > 400 \text{ or } Z < 300.}$ $\Rightarrow$ $|Z - 350| > 50.$

$$\mathbb{P}(\underbrace{|Z - \mathbb{E}[Z]| \geq k}) = \mathbb{P}(|Z - \mathbb{E}[Z]| \geq 51)$$

$$\leq \frac{\mathrm{Var}(Z)}{51} = \frac{100 \cdot (35/12)}{51^2} \approx \boxed{\frac{7}{60}}$$

$\mathbb{E}[Z'] = 375.$

$k = 25.$



Q: when can I divide by 2 in Chebyshev's?

$\hookrightarrow$ when my distribution is symmetric.

$\mathbb{P}(|Z - \mu| \geq 25) \leq a \Rightarrow \mathbb{P}(Z \geq \mu + 25) \leq a/2$

$\mathbb{P}(|Z - \mu| \geq 75) \leq b.$ $\quad \mathbb{P}(Z \leq \mu - 75) \leq b/2$

$+$

$\leq \frac{a+b}{2}.$

3 Working with the Law of Large Numbers

(a) A fair coin is tossed multiple times and you win a prize if there are more than 60% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.

(b) A fair coin is tossed multiple times and you win a prize if there are more than 40% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.

(c) A fair coin is tossed multiple times and you win a prize if there are between 40% and 60% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.

(d) A fair coin is tossed multiple times and you win a prize if there are exactly 50% heads. Which number of tosses would you prefer: 10 tosses or 100 tosses? Explain.

2 Vegas

On the planet Vegas, everyone carries a coin. Many people are honest and carry a fair coin (heads on one side and tails on the other), but a fraction p of them cheat and carry a trick coin with heads on both sides. You want to estimate p with the following experiment: you pick a random sample of n people and ask each one to flip his or her coin. Assume that each person is independently likely to carry a fair or a trick coin.

1. Given the results of your experiment, how should you estimate p?
   (*Hint:* Construct an (unbiased) estimator for p such that $E[\hat{p}] = p$.)

2. How many people do you need to ask to be 95% sure that your answer is off by at most 0.05?

1)

	HT	HH	
			n
$1-p$	honest	cheaters	p

X = fraction of heads observed

$= \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$ where $X_i = \begin{cases} 1 & \text{if } i\text{th flip is H} \\ 0 & \text{o/w} \end{cases}$

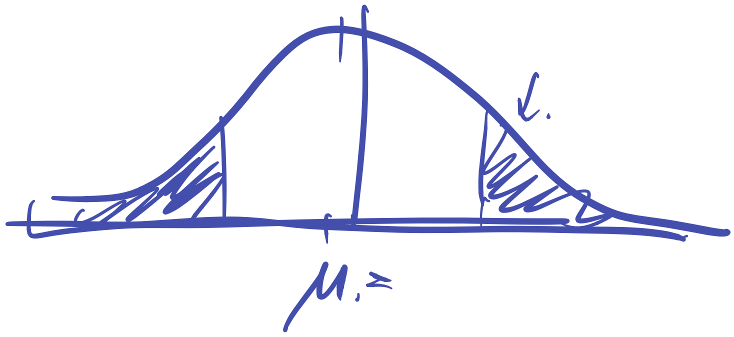
$$\mathbb{E}[X] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[X_i] = \frac{1}{n} n \cdot \mathbb{E}[X_i] = \mathbb{E}[X_i].$$

$$\mathbb{E}[X_i] = p \cdot 1 + (1-p) \cdot \frac{1}{2}$$
$$= \frac{1}{2} + \frac{p}{2} \Rightarrow p = 2 \cdot \mathbb{E}[X_i] - 1.$$

$$p = 2 \cdot \mathbb{E}[X] - 1 = \underbrace{\mathbb{E}[2X-1]}.$$

$$\Rightarrow \hat{p} = 2X - 1.$$

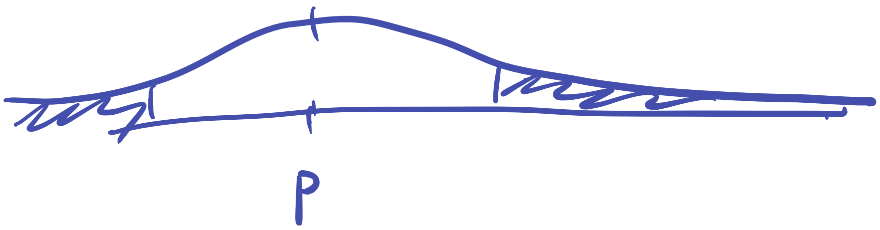
$X,$



$$\mathbb{P}(|X - \underbrace{\mathbb{E}[X]}_{\frac{p+1}{2}}| > k) \le \cdots$$

$\hat{p}$

$$\mathbb{P}(|\hat{p} - p| > k) \le \frac{Var(\hat{p})}{k^2}.$$



$$Var(\hat{p}) \cdots \qquad \underbrace{Var(X_i)} \le \frac{1}{4}$$

$$X_i = \begin{cases} 1 & \text{w/ prob } p \\ 0 & \text{o/w} \end{cases} \qquad \mathbb{E}[X_i] = p$$

$$Var(X_i) = (1-p)^2 \cdot p + (0-p)^2 \cdot (1-p)$$
$$= p(1-p)$$