

Section 9.1

Estimating tail probabilities

Basic question: R.v. is too complicated, or we have limited information about it. Can we still bound how “big” it is?

We know some of

- $\mathbb{E}X$

- $\text{Var}(X)$

- X is non-negative

We want expressions
of the form
 $P(X > t) \leq \delta.$

Ex) On an average, your ice cream shop has 1000 customers in a day. What is the probability of at least 2000 customers tomorrow?

Ans: Not enough info

Instead, we seek a bound on the probability

Let X be the number of customers on a given day. What can we say about X ?

$$EX = 1000$$

Also, $X \geq 0$.

Markov Inequality

Theorem (Markov's inequality): Let X be a non-negative r.v. (i.e., $X \geq 0$). Then, for any $t > 0$,

$$P(X \geq t) \leq \frac{EX}{t}$$

Thus, in ice cream shop example,

$$P(X \geq 2000) \leq \frac{EX}{2000} = \frac{1000}{2000} = 50\%$$

Proof of Markov's inequality

Lemma: Let X be non-negative. Then for any $t \geq 0$, we have

$$X \geq t 1_{[X \geq t]}.$$

Proof of theorem given lemma: By taking expected value of both sides of inequality in lemma, we have

$$\begin{aligned} \mathbb{E}X &\geq \mathbb{E}[t 1_{[X \geq t]}] = t \mathbb{E}[1_{[X \geq t]}] \\ &= t \cdot P(X \geq t) \end{aligned}$$

Divide both sides by t to give
Markov's $\neq$: $P(X \geq t) \leq \frac{\mathbb{E}X}{t}$. $\checkmark$

Proof of lemma

Lemma: Let X be non-negative. Then for any $t \geq 0$, we have

$$X \geq t \mathbb{1}_{[X \geq t]}.$$

There are two cases to consider:

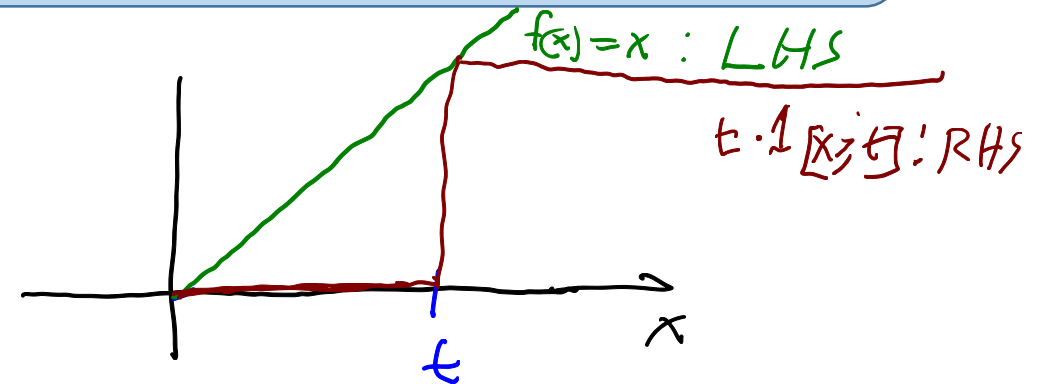
Case 1: $X < t$.

$$\text{RHS} = t \cdot \mathbb{1}_{[X \geq t]} = 0 \leq X = \text{LHS}$$

↑
since X is non-neg

Case 2: $X \geq t$

$$\text{RHS} = t \cdot \mathbb{1}_{[X \geq t]} = t \leq X = \text{LHS}$$



Ex) You roll a six-sided die until you get 100 sixes. Bound the probability that you roll at least 5000 times.

$X = \# \text{ of rolls}$, $X \geq 0$, $\mathbb{E}X = ?$

$X = X_1 + X_2 + \dots + X_{100}$ where $X_i = \# \text{ of additional rolls until rolling next } 6 \sim \text{Geom}(p_6)$

$$\mathbb{E}X = \mathbb{E}X_1 + \dots + \mathbb{E}X_{100} = 100 \cdot \mathbb{E}[\text{Geom}(p_6)] = 100 \cdot 6 = 600.$$

$$P(X \geq 5000) \leq \frac{\mathbb{E}X}{5000} \quad \text{By Markov \#}$$

$$\leq \frac{600}{5000} = \boxed{\frac{3}{25}}$$

Ex) There are 10 people in front of you in line. It takes each person $\text{Exp}\left(\frac{1}{5}\right)$ minutes to be served. Bound the probability that you wait more than 3 hours. Markov $\neq$: $P(X \geq t) \leq \frac{EX}{t}$

$$X \geq 0, \quad EX = ?$$

$$X = X_1 + X_2 + \dots + X_{10}$$

$$X_i \sim \text{Exp}\left(\frac{1}{5}\right)$$

$$EX = EX_1 + \dots + EX_{10} = 10 \cdot EX_{\text{Exp}\left(\frac{1}{5}\right)} = 10 \cdot 5$$

$$\Rightarrow \text{By Markov } \neq, \quad P(X \geq 180) \leq \frac{EX}{180} = \frac{50}{180} = \boxed{\frac{5}{18}}$$

Ex) You play the following betting game 50 times: Roll two dice, if the sum is even, you win the sum of the dice. If the sum is odd, you lose the sum of the dice. Bound the probability that you win more than \$200 in total.

X can be negative $\Rightarrow$ Markov $\neq$ does not apply

Ex) Let $X \sim \text{Exp}(1)$. Bound $P(X \geq t)$ with Markov's inequality, and compare this to the exact value of $P(X \geq t)$.

$$X \geq 0, \quad \mathbb{E}X = 1$$

$$\text{Markov } \neq: P(X \geq t) \leq \frac{\mathbb{E}X}{t} = \frac{1}{t} \quad \leftarrow \text{slow } \frac{1}{t} \text{ decay}$$

$$\text{Exact soln: } P(X \geq t) = P(\text{Exp}(1) \geq t) = e^{-t} \quad \nwarrow \text{exponential decay}$$

Strength and weakness of Markov's inequality

Strength: Very little needs to be known (or calculated) about r.v. X .
Just need $\mathbb{E}X$, and $X \geq 0$.

Weakness: Says nothing if r.v. can take negative values. The *tail bound* $\frac{\mathbb{E}X}{t}$ decays proportionally to $\frac{1}{t}$.

However, if you also have $\text{Var}(X)$ you can do better.

Chebyshev's inequality

Theorem (Chebyshev's inequality): Let X be a r.v. with $\mathbb{E}X = \mu$, $\text{Var}(X) = \sigma^2$. Then for any $t > 0$,

$$P(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2}$$

- Equivalently, $P(X \in (\mu - t, \mu + t)) \geq \frac{1 - \frac{\sigma^2}{t^2}}{1}$.
- E.g., by setting $t = 10\sigma$, we get $X \in (\mu - 10\sigma, \mu + 10\sigma)$ with probability at least $1 - \frac{\sigma^2}{(10\sigma)^2} = 1 - \frac{1}{100} = 99\%$.
- Doesn't require $X \geq 0$.
- Gives better tail bound decay $\sim \frac{1}{t^2}$.

Proof of Chebyshev's inequality

First, note: $P(|X - \mu| \geq t) = P((X - \mu)^2 \geq t^2)$

Now, use Markov's inequality on the RHS to give:

$$P((X - \mu)^2 \geq t^2) \leq \frac{\mathbb{E}[(X - \mu)^2]}{t^2} = \frac{\text{Var}(X)}{t^2} \quad \checkmark$$

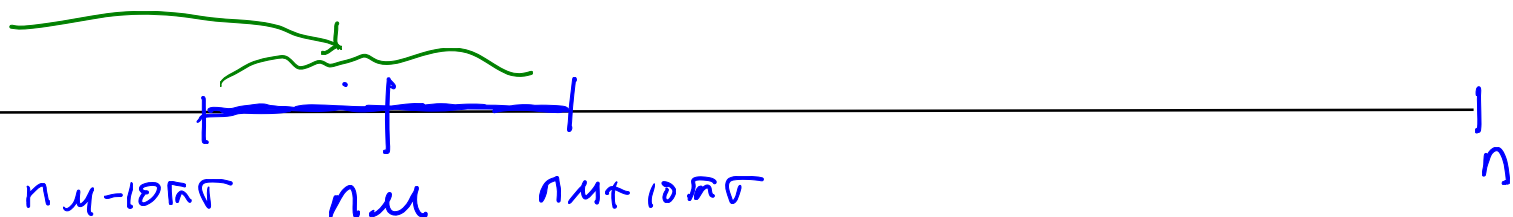
$\sqrt{n}$ deviation of a sum

Corrolary: Let $X_1, X_2 \dots X_n$ be uncorrelated r.v.s each with mean μ and variance σ^2 . Let $S_n = X_1 + X_2 + \dots + X_n$. Then

$$P(|S_n - n\mu| \geq \sqrt{nt}) \leq \frac{\sigma^2}{t^2} \quad \text{or, equivalently,} \quad P\left(\left|\frac{S_n}{n} - \mu\right| \geq \frac{t}{\sqrt{n}}\right) \leq \frac{\sigma^2}{t^2}.$$

- Proof: Chebychev gives $P(|S_n - n\mu| \geq \sqrt{nt}) \leq \frac{\text{Var}(S_n)}{(\sqrt{nt})^2} = \frac{n\sigma^2}{nt^2}$ ✓
- E.g., $S_n \in (n\mu - 10\sqrt{n}\sigma, n\mu + 10\sqrt{n}\sigma)$ with probability at least 99%.

*S_n is prob.
here*



Ex) You flip a coin 10,000 times. Let X be number of heads. Use Markov's inequality and Chebyshev's inequality to bound

$$P(X \geq 7500). \quad X \sim \text{Bin}(10,000, 1/2), \quad X \geq 0$$

$$\mathbb{E} X = 10,000 \cdot 1/2 = 5000, \quad \text{Var}(X) = 10,000 \cdot 1/2 \cdot 1/2 = 2500$$

Markov $\neq$:

$$P(X \geq 7500) \leq \frac{\mathbb{E} X}{7500} = \frac{5000}{7500} = \boxed{\frac{2}{3}}$$

Chebyshev $\neq$:

$$\begin{aligned} P(X \geq 7500) &= P(X - 5000 \geq 2500) \leq P(|X - 5000| \geq 2500) \\ &\leq \frac{\text{Var}(X)}{(2500)^2} = \frac{2500}{(2500)^2} = \boxed{\frac{1}{2500}} \end{aligned}$$

Ex) You play the following betting game 50 times: Roll two dice, if the sum is even, you win the sum of the dice. If the sum is odd, you lose the sum of the dice. Bound the probability that you win more than \$200 in total.

Set $X = \text{total gain/loss}$.

We need EX , $\text{Var}(X)$.

$$X = X_1 + X_2 + \dots + X_{50}$$

↑
gain/loss on
1st play

$$\begin{aligned} EX &= EX_1 + \dots + EX_{50} \\ &= 50 \cdot EX_1 \end{aligned}$$

$$\begin{aligned} \text{Var}(X) &= \text{Var}(X_1) + \dots + \text{Var}(X_{50}) \\ &= 50 \text{Var}(X_1) \end{aligned}$$

$$X_1 \stackrel{\text{dist}}{=} (\underbrace{u}_{1\text{st die}} + \underbrace{v}_{2\text{nd die}}) Z$$

$$Z = \begin{cases} 1 & \text{if sum is even} \\ -1 & \text{if " " " odd} \end{cases}$$

$$E X_1 = E[(u+v)z] = E[uz] + E[vz] = E u \cdot \cancel{E z} + E v \cdot \cancel{E z} = 0$$

$\begin{array}{c} | \quad | \\ \text{indep} \end{array}$
 Exercise

 $\begin{array}{c} | \quad | \\ \text{indep} \end{array}$
 Exercise

$\begin{array}{c} \text{0} \\ \text{0} \end{array}$
 Exercise

$$\begin{aligned}
 \text{Var}(X_1) &= \text{Var}((u+v)z) = E[(u+v)z]^2 - (E[(u+v)z])^2 \\
 &= E(u+v)^2 \cdot \cancel{z^2} = E[u^2 + v^2 + 2uv] \\
 &= E u^2 + E v^2 + 2 E[uv] \\
 &= \frac{91}{6} + \frac{91}{6} + 2 \cdot \frac{7}{2} \cdot \frac{7}{2} = \frac{91}{3} + \frac{49}{2}
 \end{aligned}$$

$\begin{array}{c} | \\ \text{indep} \end{array}$

$$\text{So, } \text{Var}(X) = 50 \cdot \left(\frac{91}{3} + \frac{49}{2} \right), \quad \mathbb{E} X = 0.$$

By Chebyshev $\neq$,

$$P(X \geq 200) \leq P(|X - 0| \geq 200)$$

$$\leq \frac{\text{Var}(X)}{200^2} = \boxed{\frac{50 \cdot \left(\frac{91}{3} + \frac{49}{2} \right)}{200^2}}$$