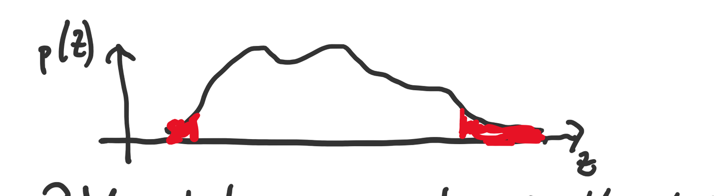
Concentration Inequalities

Thursday, October 27, 2022 3:58 PM



Goal: Provide bounds on P(a RV takes on values in the tail of the distribution)

↳ Why? Helps us with theoretical analysis & understanding of random variables & algorithms that use RVs

e.g. Sample means $Y = \frac{1}{n}\sum_{i=1}^{n} X_i$ How far is $E[Y]$ from $E[X_i]$ → ideally, "not very"

Next time: use concentration inequalities to analyze multi-armed bandit algorithms (repeated decisions with feedback)

Warmup Exercise

Suppose a slot machine (payouts always ≥0) has an expected payout of \$5. Let $p = P(\text{payout} = \$100)$. Which are poss. values for p?

A. 2% ✓ B. 4% ✓ C. 5% ✓ D. 7% ✗

C:

pay	prob
100	0.05
0	0.95

$E[\text{payout}] = \$100 \cdot 0.05 + \$0 \cdot 0.95 = \$5$

B:

pay	prob
100	0.04
99	0.01
1	0.01
0	0.94

$E[\text{payout}] = \$5$

General fact

If X is nonneg. RV, $E[X]=\mu$, then $P(X \geq t) \leq \frac{\mu}{t}$

Markov's Inequality

D:

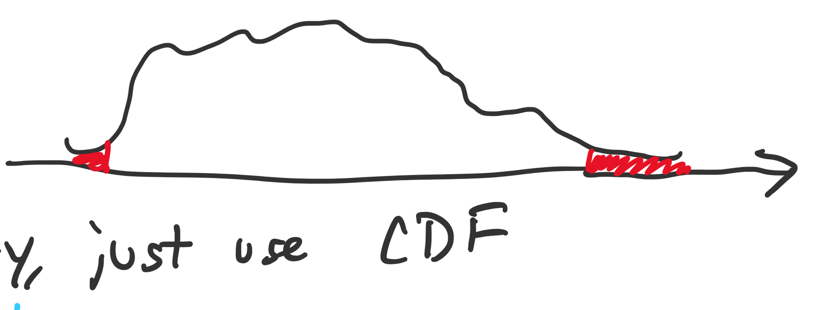
pay	prob
100	0.07
⋮	⋮

$E[\text{payout}] = \$100 \cdot 0.07 + \cdots = \$7 + \cdots$

Concentration Inequalities

Inequality	Info needed	How good is the bound?
Markov	mean	Bad
Chebyshev	mean, variance	OK
Chernoff	moment-generating function	Good
Hoeffding	must be bounded	Good

Tail Probabilities



How to compute?

- If known distribution/density, just use CDF
  e.g., Gaussian, Uniform, Exponential, etc.
- Usually:
  1) Distributions aren't known
  2) Involves combination of other RVs (e.g., sample mean, quicksort) MABs

Example:
- I have 10 biased coins
- Coin i has prob. p_i of coming up heads (unknown!)
- $p_1 + \cdots + p_{10} = 1$
- I flip all of them
- What is $P(\text{all heads})$?
  ↳ Compute an upper bound

$\sum p_i = 1$

Best bound

$P(\text{all heads}) = p_1 \cdot p_2 \cdot p_3 \cdots p_{10} \leq \boxed{\left(\frac{1}{10}\right)^{10}}$

Markov's Inequality

Let $X_i = \begin{cases} 1 & \text{if coin } i \text{ heads} \\ 0 & \text{o.w.} \end{cases}$

$Y = \sum X_i$

"all heads" $\Rightarrow Y \geq 10$

$P(Y \geq 10) \leq \frac{E[Y]}{10}$ → $E[Y] = E[\sum X_i] = \sum E[X_i] = \sum p_i = 1$

$P(Y \geq 10) \leq \boxed{\frac{1}{10}}$

Chebyshev

$P(Y \geq 10)$ $E[Y] = 1$

$var(X_i) = p_i(1-p_i) = p_i - p_i^2 \leq p_i$

$var(Y) = var(\sum X_i) = \sum var(X_i) \leq \sum p_i = 1$

$P(Y \geq 10) = P(Y - 1 \geq 9) \leq P(|Y-1| \geq 9) \leq \boxed{\frac{1}{9^2}}$

Chernoff ($\lambda = 1$)

$P(Y \geq 10)$

$P(e^{\lambda Y} \geq e^{10\lambda}) \leq \frac{E[e^{\lambda Y}]}{e^{10\lambda}} \overset{\lambda=1}{\leq} \frac{e^{e-1}}{e^{10}} = \boxed{\frac{1}{10^{3.6}}}$

$\frac{e^{e^{\lambda}-1}}{e^{10\lambda}} = \boxed{\frac{1}{10^{6.1}}}$

Fact: $1 + x \leq e^x$

$\Rightarrow E[e^{Y}] = \prod E[e^{X_i}] \Rightarrow E[e^{X_i}] = (1-p_i)e^0 + p_i e^1 = 1 + (e-1)p_i \leq e^{p_i(e-1)}$

$\leq \prod e^{p_i(e-1)} = e^{(e-1)\sum p_i} = e^{e-1}$

Chebyshev's Inequality

Markov

$P(Z \geq t) \leq \frac{E[Z]}{t}$, if Z nonneg.

Idea: use mean <u>and</u> variance

Suppose $E[X] = \mu$, $var(X) = \sigma^2$

Find a bound on $P(|X-\mu| \geq t\sigma)$ "X is more than t SDs away from its mean"

$Z = (X-\mu)^2$, $E[Z] = E[(X-\mu)^2] = var(X) = \sigma^2$

$P(|X-\mu| \geq t\sigma) = P(Z \geq t^2\sigma^2) \leq \frac{E[Z]}{t^2\sigma^2}$

$P(|X-\mu| \geq t\sigma) \leq \frac{1}{t^2}$

Chernoff Bound

Markov: If Z nonneg, $P(Z \geq t) \leq \frac{E[Z]}{t}$

- Applying Markov to $Z = (X-\mu)^2$ was a great idea
- How about $(X-\mu)^3$?
- How about $(X-\mu)^4$?

  ⋮
- How about $e^{\lambda X}$?

$\lambda \geq 0$

$$\boxed{P(X \geq t) = P(e^{\lambda X} \geq e^{\lambda t}) \leq \frac{E[e^{\lambda X}]}{e^{\lambda t}}}$$

Taylor Series

$e^{\lambda X} = 1 + \lambda X + \frac{\lambda^2}{2!}X^2 + \frac{\lambda^3}{3!}X^3 + \cdots$

$m_X(\lambda)$ $E[e^{\lambda X}] = 1 + \lambda E[X] + \frac{\lambda^2}{2!}E[X^2] + \cdots$

Moment generating function $m_X(\lambda)$

$Y = \sum X_i$, X_i indep

$m_Y(\lambda) = E[e^{\lambda Y}] = E[e^{\lambda X_1 + \lambda X_2 + \cdots + \lambda X_n}] = E[\prod e^{\lambda X_i}] \overset{\text{by indep}}{=} \prod E[e^{\lambda X_i}] = \prod_{i=1}^{n} m_{X_i}(\lambda)$

Hoeffding's Lemma

Suppose X is bounded between a and b with mean μ. Then

$m_X(\lambda) \leq \exp\left(\frac{(b-a)^2}{8}\lambda^2 + \mu\lambda\right)$

MGF for X: $E[e^{\lambda X}]$

$P(X \geq t) = P(e^{\lambda X} \geq e^{\lambda t}) \leq \frac{\exp(\quad)}{\exp(\lambda t)} = \exp\left(\frac{(b-a)^2}{8}\lambda^2 - \lambda t\right)$

Hoeffding's Inequality

Let $X_1, \ldots, X_n$ be independent bounded RVs between a, b. (not necessarily identically distributed!)

Let $Y = \frac{1}{n}\sum_{i=1}^{n}(X_i - E[X_i])$.

Then $P(Y \geq t) \leq \exp\left(-\frac{2nt^2}{(a-b)^2}\right)$

$P(Y \leq t) \leq \exp\left(-\frac{2nt^2}{(a-b)^2}\right)$