October 28

Wednesday, October 28, 2020 3:56 PM

Comp Sci 212

1. Chebyshev's Inequality

2. Chernoff bounds

Thm. Markov's Inequality

If R random variable, non-negative, $x > 0$

$$\Pr[R \geq x] \leq \frac{\mathbb{E}[R]}{x}$$

$S = \{H, T\}^{100}$ Uniform dist.

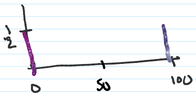
Sticky Coin - all heads / all tails w.p. $\frac{1}{2}$

$R: S \to \mathbb{R}$, $R(s) = \#$ of heads

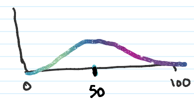
$\Pr[R \geq 75]$?

$\mathbb{E}[R] = 50$ in both cases

Sticky, $\frac{1}{2}$



Uniform



How to get better bounds for uniform dist?

Instead of looking at $\mathbb{E}[R]$, $\mathbb{E}[R^2]$.

Sticky - $\mathbb{E}[R^2] = 5{,}000$

Uniform - $R = \sum_{i=1}^{100} \mathbb{1}_{E_i}$, E_i event i^{th} coin is heads.

$$\mathbb{E}[R^2] = \mathbb{E}\left[\left(\sum_{i=1}^{100} \mathbb{1}_{E_i}\right)^2\right] \overset{\text{distributive}}{=} \mathbb{E}\left[\sum_{i,j=1}^{n} \mathbb{1}_{E_i} \mathbb{1}_{E_j}\right]$$

$$= \sum_{i,j \in [100]^2} \mathbb{E}[\mathbb{1}_{E_i} \mathbb{1}_{E_j}]$$ (linearity of expectation)

$i = j$: $\mathbb{E}[\mathbb{1}_{E_i}^2] = \mathbb{E}[\mathbb{1}_{E_i}] = \frac{1}{2}$

$i \neq j$: $\mathbb{E}[\mathbb{1}_{E_i} \mathbb{1}_{E_j}] = \mathbb{E}[\mathbb{1}_{E_i \cap E_j}] = \frac{1}{4}$

$\mathbb{1}_{E_i} \mathbb{1}_{E_j}(s) = 1$ if $s \in E_i$ and E_j, 0 o.w.

$$= \sum_{\substack{i,j \in [100]^2 \\ i = j}} \frac{1}{2} + \sum_{\substack{i,j \in [100]^2 \\ i \neq j}} \frac{1}{4} = 100 \cdot \frac{1}{2} + 9{,}900 \cdot \frac{1}{4} = 2{,}525$$

100 terms, 9,900 → generalized product rule

10,000 terms total

$(\mathbb{1}_{E_1} + \mathbb{1}_{E_2} \ldots)(\mathbb{1}_{E_1} + \mathbb{1}_{E_2} \ldots)$

$R^2: S \to \mathbb{R}$, $R^2(s) = (R(s))^2$

$\mathbb{1}_{E_1} \mathbb{1}_{E_2}$ $\mathbb{1}_{E_2} \mathbb{1}_{E_1}$

Chebyshev R random variable, positive real x

$$\Pr[|R - \mathbb{E}[R]| \geq x] \leq \frac{\mathbb{E}[R^2] - \mathbb{E}[R]^2}{x^2}$$

Proof. Let $Y = (R - \mathbb{E}[R])^2$

$\Pr[|R - \mathbb{E}[R]| \geq x] = \Pr[Y \geq x^2] \leq \frac{\mathbb{E}[Y]}{x^2}$ (Markov's inequality)

$Y \geq x^2$ iff $|R - \mathbb{E}[R]| \geq x$

$\mathbb{E}[Y] = \mathbb{E}[R^2] - \mathbb{E}[R]^2$

$\mathbb{E}[(R - \mathbb{E}[R])^2] = \mathbb{E}[R^2 - 2R\mathbb{E}[R] + \mathbb{E}[R]^2]$

$= \mathbb{E}[R^2] - \mathbb{E}[2R\mathbb{E}[R]] + \mathbb{E}[R]^2$ (linearity of expectation)

$= \mathbb{E}[R^2] - 2\mathbb{E}[R]\mathbb{E}[R] + \mathbb{E}[R]^2$

$= \mathbb{E}[R^2] - 2\mathbb{E}[R]^2 + \mathbb{E}[R]^2$

$= \mathbb{E}[R^2] - \mathbb{E}[R]^2$ $\square$

$\mathbb{E}[a \cdot R] = a\mathbb{E}[R]$, a is number

Def. Variance $\mathrm{Var}[R] = \mathbb{E}[R^2] - \mathbb{E}[R]^2$

Sticky $\mathrm{Var}[R] = 5{,}000 - 2500 = 2{,}500$

Uniform $\mathrm{Var}[R] = 2{,}525 - 2500 = 25$

$\Pr[R \geq 75] \leq \Pr[|R - 50| \geq 25]$

If $R \geq 75$, then $|R - 50| \geq 25$

$\leq \frac{25}{25^2} = \frac{1}{25}$ uniform

$\leq \frac{2500}{25^2} = 4$ sticky

Same if $\Pr[R \leq 25]$

If $R \leq 25$, then $|R - 50| \geq 25$

Variance is a measure of how spread out a distribution is.

$\mathrm{Var} = \text{std. dev}^2$

Thm. If R_1, R_2 are independent random variable,

$$\mathrm{Var}[R_1 + R_2] = \mathrm{Var}[R_1] + \mathrm{Var}[R_2]$$

Proof. $\mathrm{Var}[R_1 + R_2] = \mathbb{E}[(R_1 + R_2)^2] - (\mathbb{E}[R_1 + R_2])^2$

$= \mathbb{E}[R_1^2] + 2\mathbb{E}[R_1 R_2] + \mathbb{E}[R_2^2] - \mathbb{E}[R_1]^2 - 2\mathbb{E}[R_1]\mathbb{E}[R_2] - \mathbb{E}[R_2]^2$

$\mathbb{E}[R_1 + R_2] = \mathbb{E}[R_1] + \mathbb{E}[R_2]$; $\mathbb{E}[R_1 R_2] = \mathbb{E}[R_1]\mathbb{E}[R_2]$ by independence

$= \mathrm{Var}[R_1] + \mathrm{Var}[R_2]$ $\square$