M362K: March 27th, 2024.

More on Expectation.

Recall. Let X be a discrete random variable w/ the probability mass function p_X.
The **expectation** of X is defined as

$$\boxed{\mathbb{E}[X] = \sum_{\text{all } x} x \cdot p_X(x)}$$ if the sum exists

Proposition. Let X be a random variable w/ the support $\{x_1, \dots, x_m\}$.
Let g be a real-valued function well-defined on the support of X.
Then,

$$\boxed{\mathbb{E}[g(X)] = \sum_{i=1}^{m} g(x_i) p_X(x_i)}.$$

Problem. Let X be a r.v. corresponding to the result of rolling an ordinary fair die.

Define $Y = |X - 2|$.

Find $\mathbb{E}[Y]$.

→: Method I. By def'n of $\mathbb{E}$.

x	y	prob.
1	1	1/6
2	0	1/6
3	1	1/6
4	2	1/6
5	3	1/6
6	4	1/6

y	0	1	2	3	4
$p_Y(y)$	1/6	1/3	1/6	1/6	1/6

$$\mathbb{E}[Y] = 0 \cdot \frac{1}{6} + 1 \cdot \frac{1}{3} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6}$$
$$= \frac{11}{6}$$

Method II. Proposition.

$$\mathbb{E}[Y] = |1-2| \cdot \frac{1}{6} + |2-2| \cdot \frac{1}{6} + |3-2| \cdot \frac{1}{6} + |4-2| \cdot \frac{1}{6} + |5-2| \cdot \frac{1}{6} + |6-2| \cdot \frac{1}{6} = \frac{11}{6}$$

□

Def'n. For $k=1,2,\ldots,$ $\boxed{\mathbb{E}[X^k]}$ is called the k^{th} raw moment of the r.v. X.

Note. The 1^{st} raw moment is the expectation.

Def'n. For a random variable X w/ mean $\mu_X = \mathbb{E}[X]$, its variance is defined as

$$\boxed{\mathrm{Var}[X] = \mathbb{E}\left[(X-\mu_X)^2\right]}$$

Remark.

$$\begin{aligned}\mathrm{Var}[X] &= \mathbb{E}\left[(X-\mu_X)^2\right]\\ &= \mathbb{E}\left[X^2 - 2\cdot\mu_X\cdot X + \mu_X^2\right] \quad \text{linearity}\\ &= \underbrace{\mathbb{E}[X^2] - 2\cdot\mu_X\cdot\underbrace{\mathbb{E}[X]}_{\mu_X}}_{-2\cdot\mu_X^2} + \mu_X^2\end{aligned}$$

$$\boxed{\mathrm{Var}[X] = \mathbb{E}[X^2] - \left(\mathbb{E}[X]\right)^2}$$

Markov's Inequality.

Let $X \geq 0$ be a r.v. and let $a > 0$.

Define the the event $\boxed{A = \{X \geq a\}}$ ← the tail event

Introduce the indicator r.v. $\mathbb{I}_A$.

Then, $\mathbb{E}[X] = \mathbb{E}[X \cdot \underbrace{1}_{\mathbb{I}_A + \mathbb{I}_{A^c}}]$

$$= \mathbb{E}[X \cdot \mathbb{I}_A + X \cdot \mathbb{I}_{A^c}] =$$

$$= \mathbb{E}[X \cdot \mathbb{I}_A] + \underbrace{\mathbb{E}[\underbrace{X}_{\geq 0} \underbrace{\mathbb{I}_{A^c}}_{\geq 0}]}_{\geq 0}$$

linearity of $\mathbb{E}$

≥ 0

$$\mathbb{E}[X] \geq \mathbb{E}[X \cdot \mathbb{I}_A] \geq \mathbb{E}[a \, \mathbb{I}_A] = a \cdot \underbrace{\mathbb{E}[\mathbb{I}_A]}_{\mathbb{P}[A]}$$

on A, we know that $X \geq a$

$$\mathbb{E}[X] \geq a \cdot \mathbb{P}[A] \implies \boxed{\mathbb{P}[X \geq a] \leq \frac{\mathbb{E}[X]}{a}}$$

Markov's Inequality ✓

Example. From past experience, a professor knows that the test score of a student taking his final is a r.v. X w/ mean $\mu_X = 75$.

Then, an upper bound on the probability that a student's test score is @ least 85 is

$$\mathbb{P}[X \geq 85] \leq \frac{\mathbb{E}[X]}{85} = \frac{75}{85} = \frac{15}{17}.$$

Chebyshev's Inequality.

... is a corollary of Markov's inequality.

Let Y be a r.v. w/ a finite mean μ_Y and variance σ_Y^2.
Let $k > 0$.

$$\boxed{\mathbb{P}[|Y - \mu_Y| > k] \leq \frac{\sigma_Y^2}{k^2}}$$

Proposition. If $Var[Y] = 0$, then $Y = \mu_Y$ (w/ probability 1).