Mean - Variance

June 18, 2023

Motivation

- The actual value of a RV is ... unpredictable
- Single number to summarize information of RV
- allow to compare between some random variables

Expectation - Variance

- **Expectation** (mean, average value) is a measure of the **center** or middle of the probability distribution, considered as **representative value** of RV
- **Variance** is a measure of the **dispersion**, or **variability** in the distribution

Motivation example

- You are offered 2 games
- Throw a fair dice
- **Game A**: if the result is an even integer, you win \$20; otherwise, you lose \$10
- **Game B**: you either win \$15 if the outcome of the die is 5 or 6; otherwise, you lose \$5

If you had to choose between games A and B, which one would you prefer?

On average, which game has better profit? Need to compare expected profit of each game

If you had to choose between games A and B, which one would you prefer?

On average, which game has better profit? Need to compare expected profit of each game

Suppose we play game A 1000 times. Since there is an equal chance to win or lose, we expected to win 500 times and lose 500 times.

$$\text{total profit} = 500 \times 20 + 500 \times (-10)$$

Profit per game A

$$\frac{500 \times 20 + 500 \times (-10)}{1000}$$

$$= \frac{1}{2} \times 20 + \frac{1}{2} \times (-10) = \$5$$

Similar argument leads profit per game B

$$\frac{1}{3} \times 15 + \frac{2}{3} \times (-5) = \$1.66667$$

The gain here is smaller than that of Game A and, unless someone cannot afford the loss of $10, they should choose the first game.

Similar argument leads profit per game B

$$\frac{1}{3} \times 15 + \frac{2}{3} \times (-5) = \$1.66667$$

The gain here is smaller than that of Game A and, unless someone cannot afford the loss of $10, they should choose the first game.

Expected profit = weighted average of losing and winning with probability as the respective weight

$$m_1P(X = m_1) + \cdots + m_kP(X = m_k)$$

Expected profit = weighted average of losing and winning with probability as the respective weight

$$m_1 P(X = m_1) + \cdots + m_k P(X = m_k)$$

Table of contents

1. Expectation
2. Variance

$E(X)$ or μ_X- Expectation of random variable X

- **Discrete case**: $Range(X)$ is countable

$$E(X) = \mu_X = \sum_{x \in Range(X)} xP(X = x)$$

- **Continuous case**: $Range(X)$ is uncountable

$$E(X) = \mu_X = \int_{-\infty}^{\infty} x f_X(x) dx \text{ where } f_X \text{ is pdf of } X$$

$E(X)$ **is a "representative" value of** X, also called by expected value, average value or mean

Example

Diane has an arts and crafts internet homebusiness.A few months after she had started her business, she estimated that the number of items she sells per day ranges from one to five, with the respective probabilities as follows.

Number of items sold	1	2	3	4	5
Probability	0.25	0.15	0.35	0.15	0.10

The expected number of items she sells during a day is

$$1(0.25)+2(0.15)+3(0.35)+4(0.15)+5(0.10)$$

Example

A lot containing 7 components is sampled by a quality inspector; the lot contains 4 good components and 3 defective components. A sample of 3 is taken by the inspector. Find the expected value of the number of good components in this sample.

Solution

- pmf

$$p(x) = \frac{\binom{4}{x}\binom{3}{3-x}}{\binom{7}{3}}$$

x	0	1	2	3
p	$1/35$	$12/35$	$18/35$	$4/35$

- Expected number of good components

$$0 \times \frac{1}{35} + 1 \times \frac{12}{35} + 2 \times \frac{18}{35} + 3 \times \frac{4}{35} = 1.7$$

Example

Let X be the random variable that denotes the life in hours of a certain electronic device. The probability density function is

$$f(x) = \begin{cases} \frac{20,000}{x^3} & \text{if } x > 100 \\ 0 & elsewhere \end{cases}$$

Find the expected life of this type of device

Solution

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$
$$= \int_{100}^{\infty} x \times \frac{20,000}{x^3}dx = 200$$

(hours)

Practice

The number of notebooks sold during a week is described by a random variable X with pmf

$$P(X = x) = \frac{2x + 3}{24}, x = 0, 1, 2, 3$$

The store orders every week from its supplier 3 notebooks at a price of \$250 each, under the following agreement: the new notebooks arrive at the store on Monday morning and any notebook not sold during the week can be returned to the supplier at the price of \$210. If the store sells notebooks at a price of \$325, **find the store's expected profit during a week**

Solution

Let Y be profit during a week then

$$Y = 325X + (3 - X)210 - 3(250)$$

Sold number x	0	1	2	3
Profit y	-120	-5	110	225
Prob	3/24	5/24	7/24	9/24

$$E(Y) = \frac{3}{24}(-120) + \frac{5}{24}(-5) + \frac{7}{24}(110) + \frac{9}{24}(225)$$

Solution

Let Y be profit during a week then

$$Y = 325X + (3 - X)210 - 3(250)$$

Sold number x	0	1	2	3
Profit y	-120	-5	110	225
Prob	3/24	5/24	7/24	9/24

$$E(Y) = \frac{3}{24}(-120) + \frac{5}{24}(-5) + \frac{7}{24}(110) + \frac{9}{24}(225)$$

Expectation of a function of RV

$$E(h(X)) = \begin{cases} \sum_x h(x)P(X = x) & X \text{ is a discrete RV} \\ \int_{-\infty}^{\infty} h(x) \underbrace{f_X(x)}_{\text{pdf of } X} \, dx & X \text{ is a continuous RV} \end{cases}$$

Example

Let X be a random variable with density function

$$f(x) = \begin{cases} \frac{x^2}{3} & \text{if } 1 < x < 2 \\ 0 & elsewhere \end{cases},$$

Find the expected value of $g(X) = 4X + 3$.

Solution

$$\begin{aligned} E(g(X)) &= \int_{-\infty}^{\infty} g(x)f(x)dx \\ &= \int_{1}^{2} (4x+3)\frac{x^2}{3}dx \\ &= 8 \end{aligned}$$

In general

If X is a continuous random variable and $g(x) = ax + b$ then

$$E(g(X)) = \int_{-\infty}^{\infty} (ax+b) \underbrace{f_X(x)}_{\text{p.d.f of } X} dx$$

$$= a \int_{-\infty}^{\infty} x f_X(x) dx + b \int_{-\infty}^{\infty} f_X(x) dx$$

$$= aE(X) + b$$

If X is a discrete random variable and $g(X) = aX + b$ then

$$E(aX + b) = \sum_x (ax + b)P(X = x)$$

$$= a\underbrace{\sum_x xP(X = x)}_{E(X)} + b\underbrace{\sum_x P(X = x)}_{1}$$

$$= aE(X) + b$$

Linear property of expectation

- Expectation of linear combination

$$E(aX+b) = aE(X)+b$$

- Expectation of sum

$$E(g(X)+h(X)) = E(g(X)) + E(h(X))$$

$$E(X_1+\cdots+X_n) = E(X_1)+\cdots+E(X_n)$$

Example

I offer you to let you play a game where you pay a \$20 entrance fee, and then I let you roll a fair 6-sided die, and pay you the rolled value times \$5. What is your expected change in money

Solution

- X= rolled value and Y= your gain
- $Y = 5X - 20$
- pmf of X

x	1	2	3	4	5	6
$P(X = x)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

- $EX = \frac{1}{6}(1) + \cdots + \frac{1}{6}(6) = 7/2$
- $EY = 5EX - 20 = -5/2$

Solution

- X= rolled value and Y= your gain
- $Y = 5X - 20$
- pmf of X

x	1	2	3	4	5	6
$P(X = x)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

- $EX = \frac{1}{6}(1) + \cdots + \frac{1}{6}(6) = 7/2$
- $EY = 5EX - 20 = -5/2$

Example

A construction firm has recently sent in bids for 3 jobs worth (in profits) 10, 20, and 40 (thousand) dollars. If its probabilities of winning the jobs are respectively .2, .8, and .3, what is the firm's expected total profit?

Solution

- X_i: profit from job i
- Total profit $X = X_1 + X_2 + X_3$
- Expected total profit

$$E(X) = E(X_1) + E(X_2) + E(X_3)$$

with

- $E(X_1) = 10.(0.2) = 2$
- $E(X_2) = 20.(0.8) = 16$
- $E(X_3) = 40.(0.3) = 12$

Table of contents

1. Expectation
2. Variance

Dispersion or variability

- Throw a fair dice
- Game A: win \$20 if the outcome is an even integer, lose \$10 otherwise.
- Game B: win \$60 if the outcome is an even integer, lose \$50 otherwise.
- Average profit from both game is \$5 but the second game appears to be *more risky*

Variance of X is

$$\sigma_X^2 = \operatorname{Var}(X) = E(X - EX)^2$$

standard deviation is the square root of variance

$$\sigma_X = \sqrt{\operatorname{Var}(X)}$$

Standard deviation has the same unit as X

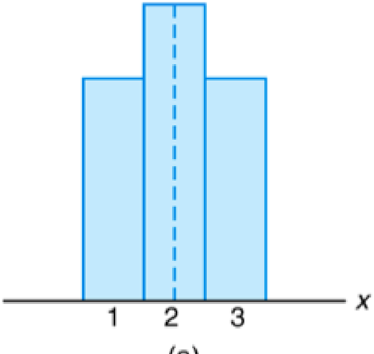
Meaning

- to present how the values of X "spread" around μ_X
- Are the other values of X usually close to μ_X or can be far away?

Meaning

- to present how the values of X ”spread” around μ_X
- Are the other values of X usually close to μ_X or can be far away?

Distribution on the right has greater variance



Property

- $$\text{Var}(X) = E(X^2) - (E(X))^2$$

  where

  $$E(X^2) = \begin{cases} \sum x^2 P(X = x), & \text{discrete RV} \\ \int_{-\infty}^{\infty} x^2 f_X(x)dx, & \text{continuous RV} \end{cases}$$

- $$Var(aX + b) = a^2 Var(X)$$

Proof

Thanks to linear property of expectation

$$
\begin{aligned}
\mathrm{Var}(X) &= E[(X-\mu)^2] \\
&= E[X^2 - 2\mu X + \mu^2] \\
&= E[X^2] - E[2\mu X] + E[\mu^2] \\
&= E[X^2] - 2\mu E[X] + \mu^2 \\
&= E[X^2] - \mu^2
\end{aligned}
$$

Example

Compute $Var(X)$ when X represents the outcome when we roll a fair die.

$$E(X) = \frac{1}{6}(1 + \cdots + 6) = \frac{21}{6}$$

$$E(X^2) = \frac{1}{6}(1^2 + \cdots + 6^2) = \frac{91}{6}$$

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \frac{35}{12}$$

Example

Compute $Var(X)$ when X represents the outcome when we roll a fair die.

$$E(X) = \frac{1}{6}(1 + \cdots + 6) = \frac{21}{6}$$

$$E(X^2) = \frac{1}{6}(1^2 + \cdots + 6^2) = \frac{91}{6}$$

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \frac{35}{12}$$

Practice

If the weather is good (which happens with probability 0.6). Alice walks the 2 miles to class at a speed of $V = 5$ miles per hour, and otherwise rides her motorcycle at a speed of $V = 30$ miles per hour. What is the mean and variance of the speech V to get to class?