Expectation - Variance

April 24, 2021

Motivation

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

- **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

- spin a wheel of fortune many time
- at each spin, $m_1, m_2, ..., m_n$ comes up with corresponding probability $p_1, p_2, ..., p_n$, monetary reward from spin
- What is the amount of money that you “expect” to get “per spin”?

- spin k times
- m_i appears k_i times
- total amount received $k_1 m_1 + k_2 m_2 + \ldots + k_n m_n$
- amount received per spin

$$M = \frac{k_1 m_1 + k_2 m_2 + \ldots + k_n m_n}{k}$$

- k is large

$$\frac{k_i}{k} \approx p_i = P(m_i)$$

-

$$M \approx p_1 m_1 + p_2 m_2 + ... + p_n m_n$$

Expectation (mean or average value)

$$\mu_X = E(X) = \begin{cases} \displaystyle\sum_{x \in Range(X)} x\, p_X(x) & \text{for discrete RV} \\ \int_{-\infty}^{\infty} x\, f_X(x)\, dx & \text{for continuous RV} \end{cases}$$

$E(X)$ is a ”representative” value of X

Meaning

- $E(X)$ is the weighted average value of X
- Think of X as a quantity defined by outcome of an experiment
- then if we repeat the experiment many times, the average value of X over all times will be about $E(X)$

Meaning

- $E(X)$ is the weighted average value of X
- Think of X as a quantity defined by outcome of an experiment
- then if we repeat the experiment many times, the average value of X over all times will be about $E(X)$

Meaning

- $E(X)$ is the weighted average value of X
- Think of X as a quantity defined by outcome of an experiment
- then if we repeat the experiment many times, the average value of X over all times will be about $E(X)$

Consider two independent coin tosses, each with a 3/4 probability of a head, and let X be the number of heads obtained.
pmf of X

x	0	1	2
$P(X=x)$	$\left(\frac{1}{4}\right)^2$	$2.\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)$	$\left(\frac{3}{4}\right)^2$
$xP(X=x)$	$0.\left(\frac{1}{16}\right)$	$1.\left(\frac{6}{16}\right)$	$2.\left(\frac{9}{16}\right)$

Expectation of X

$$EX = 0.\left(\frac{1}{16}\right) + 1.\left(\frac{6}{16}\right) + 2.\left(\frac{9}{16}\right) = \frac{24}{16} = \frac{3}{2}$$

Find $E(X)$ where X is the outcome when we roll a fair dice

Solution

pmf of X:

$$P(X = i) = \frac{1}{6}, \quad i = 1, \ldots, 6$$

$$E(X) = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + \cdots + 6\left(\frac{1}{6}\right) = \frac{7}{2}$$

Let I be the indicator variable of event A

$$I = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A^c \text{ occurs} \end{cases}$$

$$E(I) = 1p(1) + 0p(0) = P(A)$$

Let I be the indicator variable of event A

$$I = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A^c \text{ occurs} \end{cases}$$

$$E(I) = 1p(1) + 0p(0) = P(A)$$

Practice

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

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)

Compute $E(X^2)$ if

x	0	1	2
$P(X=x)$	$\frac{1}{16}$	$\frac{6}{16}$	$\frac{9}{16}$

$$E(h(X)) = \begin{cases} \sum_{x} h(x) p_X(x) & discreteRV \\ \int_{-\infty}^{\infty} h(x) f_X(x) dx & continuousRV \end{cases}$$

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

x	1	2	3	4	5	6
$P(X = x)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$
y	-15	-10	-5	0	5	10
$P(Y = y)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

- $EY = -5/2$

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

x	1	2	3	4	5	6
$P(X = x)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$
y	-15	-10	-5	0	5	10
$P(Y = y)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

- $EY = -5/2$

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

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

Linear property of expectation

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

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

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

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

- $EX = 7/2$, $EY = 5EX - 20 = -5/2$

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

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

- $EX = 7/2$, $EY = 5EX - 20 = -5/2$

Variance of X is

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

standard deviation is the square root of variance

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

Standard deviation has the same unit as X

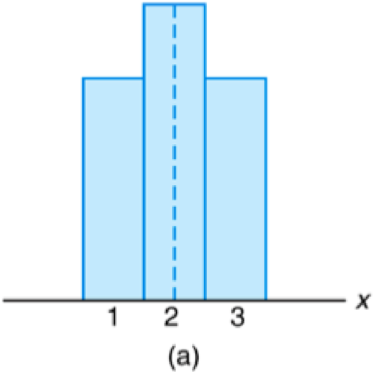
Meaning

- 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?

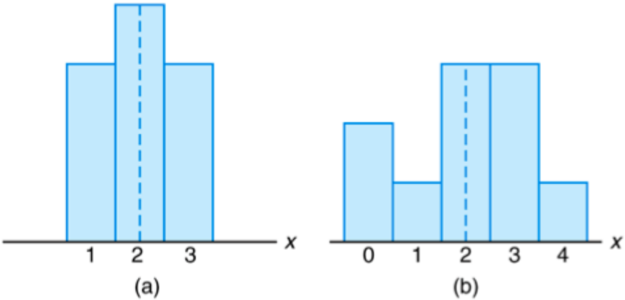
Meaning

- 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



the probability that the random variable assumes a value within a certain interval about the mean is greater than for a similar random variable with a larger standard deviation.



The values of a random variable more concentrate about the mean than for a similar random variable with a larger standard deviation.

Property

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

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

Proof for discrete random variable

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

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

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

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

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

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

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

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

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

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

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?

3 RV with same mean but very far away

$$W = 0 \quad \text{with probability } 1$$

$$Y = \begin{cases} -1 & \text{with probability } \frac{1}{2} \\ +1 & \text{with probability } \frac{1}{2} \end{cases}$$

$$Z = \begin{cases} -100 & \text{with probability } \frac{1}{2} \\ +100 & \text{with probability } \frac{1}{2} \end{cases}$$

Practice

Two balls are chosen randomly from an urn containing 8 white, 4 black, and 2 orange balls. Suppose that we win \$2 for each black ball selected and we lose \$1 for each white ball selected. Let X denote our winnings. Find the pmf of X, and calculate $E(X)$, $\text{Var}(X)$.

Use expected value to making decision

Example - Quiz problem

A quiz consists of 2 questions.

- Q1 will be answered correctly with prob 0.8 and you receive a reward of \$100
- Q2 will be answered correctly with prob 0.5 and you receive a reward of \$200
- If the first question attempted is answered incorrectly, the quiz terminates
- If the first question attempted is answered correctly, you are allowed to attempt the second question.

Which one to answer first to maximize expected value of total prize?

Example - Quiz problem

A quiz consists of 2 questions.

- Q1 will be answered correctly with prob 0.8 and you receive a reward of \$100
- Q2 will be answered correctly with prob 0.5 and you receive a reward of \$200
- If the first question attempted is answered incorrectly, the quiz terminates
- If the first question attempted is answered correctly, you are allowed to attempt the second question.

Which one to answer first to maximize expected value of total prize?

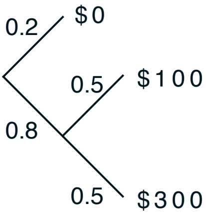
Example - Quiz problem

A quiz consists of 2 questions.

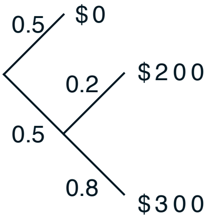
- Q1 will be answered correctly with prob 0.8 and you receive a reward of \$100
- Q2 will be answered correctly with prob 0.5 and you receive a reward of \$200
- If the first question attempted is answered incorrectly, the quiz terminates
- If the first question attempted is answered correctly, you are allowed to attempt the second question.

Which one to answer first to maximize expected value of total prize?

Solution



Question 1
Answered 1st



Question 2
Answered 1st

Let X be the value of prize if you choose to answer Q1 first

$$
\begin{aligned}
P(X=0) &= .2 \\
P(X=100) &= (.8)(.5) = .4 \\
P(X=300) &= (.8)(.5) = .4
\end{aligned}
$$

$$E(X) = 0(.2) + 100(.4) + 300(.4) = 160$$

Let Y be the value of prize if you choose to answer Q1 first

$$P(Y=0) = .5$$
$$P(Y=200) = (.5)(.2) = .1$$
$$P(X=300) = (.8)(.5) = .4$$

$$E(Y) = 0(.5) + 200(.1) + 300(.4) = 140 < E(X)$$

Q1 should be chosen to answer firstly