# 1.6 Mean and Variance of a Distribution

1.6.1 Mean of a Discrete Random Variable

1.6.2 Variance of a Discrete Random Variable

1.6.3 Mean of a Continuous Random Variable

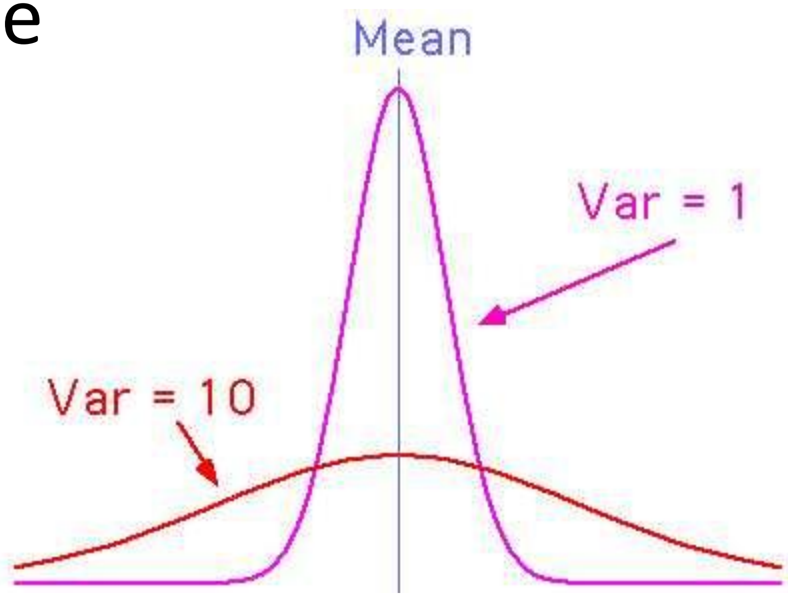
1.6.4 Variance of a Continuous Random Variable

1.6.5 Properties of Mean and Variance

1.6.6 Summary

**3 April 2018**

Giovanni Merola

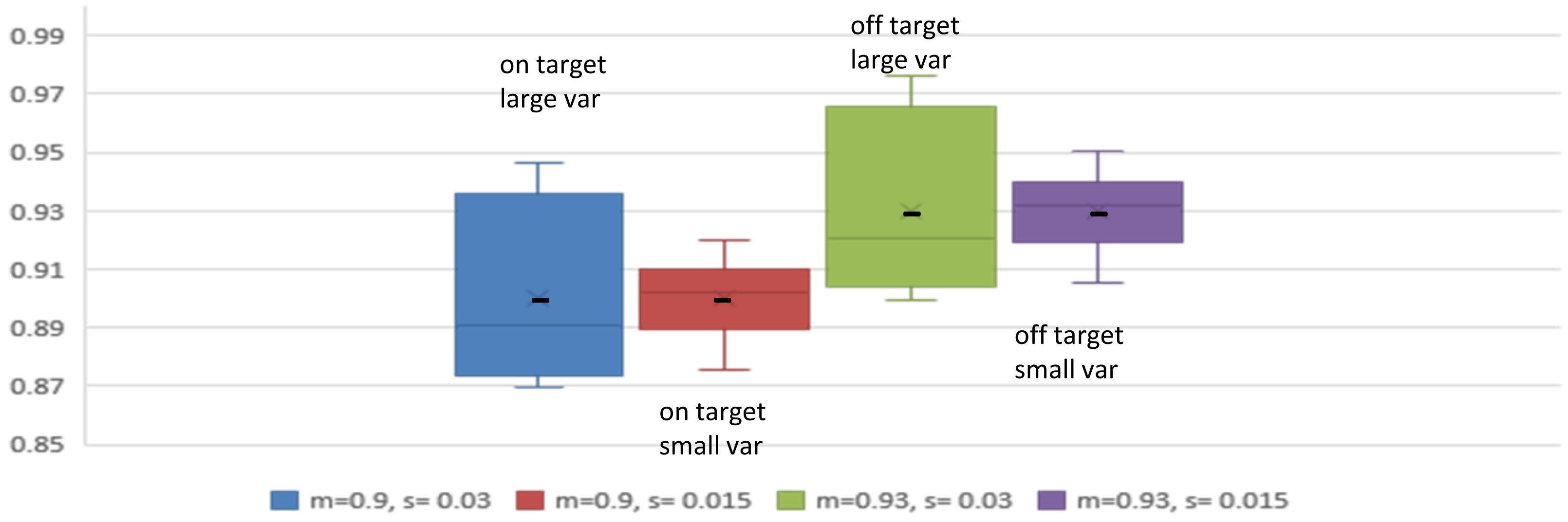


# 1.6.0 Mean and Variance

- A random variable can take many (possibly infinite) values. These are described by its distribution. We summarize a distribution with few typical values, most important are mean and variance.
- The mean is a measure of *centre* or *location*;
- The variance is a measure of *variability* (*spread*) or *scale*

# 1.6.0 Mean and Variance

- remember the chemical reactor producing a polymer example? The target density is 9.0g/cc
- Every day the density of the product is slightly different, it is a random variable depending on many factors. During the month the density is distributed according to some distribution.



# 1.6.1 Mean of a Discrete Random Variable

For a discrete random variable $X$, the *mean* (also called *expectation* or *expected value*) is

$$\mu = \mathrm{E}[X] = \sum_{j=1}^{n} x_j \, p(x_j)$$

where $p$ is the pmf of $X$ and $x_j$'s are possible values of $X$. The mean can be seen as the weighted average of all the values of $X$, i.e. $x_1, x_2, \cdots, x_n$.

# 1.6.1 Mean of a Discrete Random Variable

## Example 1

The diameters of steel spheres can take three values: 0.5mm, 0.52mm and 0.55mm, with probability 0.2, 0.5 and 0.3 respectively. Find the expected diameter of a sphere.

## Solution

Let $d$ denote the diameter of a sphere. Its expected value is equal to $E(d) = 0.5 * 0.2 + 0.52 * 0.5 + 0.55 * 0.3 = 0.525$

■

# 1.6.1 Mean of a Discrete Random Variable

## Example 2

What is the expected value of a die roll?

## Solution

The die can take values $\{1, 2, 3, 4, 5, 6\}$ with equal probability $1/6$. Therefore, the expected value is

$$E(x) = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = \frac{21}{6} = 3.5$$

This does not mean that the die takes the value 3.5 (!) but simply that taking the average of many rolls will give this value. It's a measure of centre. ■

# 1.6.1 Mean of a Discrete Random Variable: problem

The mean mass of chlorine is 35.5 unit. Chlorine atoms exist naturally in two types, one with 35 unit and another type with 37 unit. What is the fraction (proportion) of chlorine atoms that have mass of 35 unit?

[hint: the fraction is the probability]

# 1.6.1 Mean of a Discrete Random Variable

We can generalize to the following theorem.

Let $X$ be a discrete random variable and let $g(X)$ be a function of $X$. Then the expected value of $g(X)$ is

$$\mathrm{E}[g(X)] = \sum_j g(x_j)p(x_j) \quad ; \quad j = 1,2,\cdots,n$$

where $p$ is the pmf of $X$ and $x_j$'s are possible values of $X$.

This means that we can take the expectation of a function of the variable.

# 1.6.1 Mean of a Discrete Random Variable: example

**Example**

Consider $X = \{1, 2, 3, 4, 5, 6\}$ and

$$Y = g(x_i) = \begin{cases} y_1 = 1, X = 1, 2, 3, \\ y_2 = -1, X = 4, 5, 6 \end{cases}$$

$E(Y) = \sum_{i=1,2} \sum_{j=1,..,6} y_i P\left(x_j\right) = (1)\left(\frac{1}{6} + \frac{1}{6} + \frac{1}{6}\right) + (-1)\left(\frac{1}{6} + \frac{1}{6} + \frac{1}{6}\right) = 0$

We knew this by considering that $P(Y = 1) = P(Y = -1) = \frac{1}{2}$ for which

$$E(Y) = (1)\frac{1}{2} + (-1)\frac{1}{2} = 0$$

# 1.6.2 Variance of a Discrete Random Variable

The *variance* of random variable $X$ is the expected value of the function of a random variable $\left(X - \mathrm{E}(X)\right)^2$. For discrete case,

$\sigma^2 = \mathrm{Var}(X) = \mathrm{E}\left[\left(X - \mathrm{E}(X)\right)^2\right] = \sum_j \left(x_j - \mathrm{E}(X)\right)^2 p\left(x_j\right) =$
$\mathrm{E}(\mathrm{X}^2) - \mathrm{E}(X)^2$. (Last equation useful for computation)

The variance provides a measure of spread of $X$ around its mean $\mathrm{E}(X)$. The variance is always nonnegative, $\sigma^2 \geq 0$.

# 1.6.2 Variance of a Discrete Random Variable

$\underline{Proof\ of\ \sigma^2 = \mathrm{Var}(X) = \mathrm{E}(\mathrm{X}^2) - \mathrm{E}(X)^2.}$

$$\mathrm{Var}(X) = \sum_j \left(x_j - \mathrm{E}(X)\right)^2 p\left(x_j\right) =$$

$$\sum_j \left(x_j^2 - 2x_j\mathrm{E}(X) + \mathrm{E}(X)^2\right)p(x_j)$$

$$= \sum_j x_j^2 p_j - 2E(X)\sum_j x_j p\left(x_j\right) + E(X)\sum_j p(x_j)$$

$$= E(X^2) - 2E(X)^2 + E(X)^2 = \mathrm{E}(\mathrm{X}^2) - \mathrm{E}(X)^2$$

# 1.6.2 Variance of a Discrete Random Variable

The standard deviation $\sigma$ is another measure of spread defined as

$$\sigma = \sqrt{\mathrm{Var}(X)}.$$

It is easier to interpret as it has the same units as $X$.

# 1.6.2 Variance of a Discrete Random Variable

**Example 2**

The random variable $X$ describes the number of heads in a single toss of a fair coin. It has pmf

$$\begin{cases} 0.5, if\ X = 0 \\ 0.5, if\ X = 1 \\ 0, otherwise \end{cases}$$

Obtain its mean $\mathrm{E}(X)$ and variance $\mathrm{Var}(X)$.

# 1.6.2 Variance of a Discrete Random Variable

**Solution**

$$\mathrm{E}(X) = \sum_j x_j\, p(x_j) = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}$$

$$\mathrm{Var}(X) = \mathrm{E}\left[\left(X - \mathrm{E}(X)\right)^2\right] = \sum_j \Bigg(x_j - \underbrace{\mathrm{E}(X)}_{1/2}\Bigg)^2 p(x_j) = \tfrac{1}{2}\left[\left(-\tfrac{1}{2}\right)^2 + \left(\tfrac{1}{2}\right)^2\right] = \tfrac{1}{4}.$$

Or

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - E(X)^2 = \left[0\,\frac{1}{2} + 1\,\frac{1}{2}\right] - \left(\frac{1}{2}\right)^2 = \frac{1}{4} \;\blacksquare$$

# 1.6.3 Mean of a Continuous Random Variable

For a continuous random variable $X$, its *mean* is

$\mu = \mathrm{E}[X] = \int_{-\infty}^{\infty} x f(x) dx$

where $f$ is the pdf of $X$.

Generalizing, for any function $g(X)$ the expected value is

$$\mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx$$

where $f$ is the pdf of $X$.

# 1.6.4 Variance of a Continuous Random Variable

For continuous case, the variance of random variable $X$ is

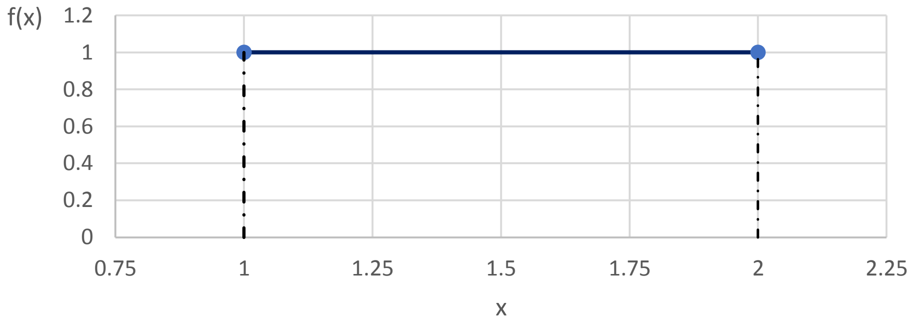
$$\sigma^2 = \mathrm{Var}(X) = \mathrm{E}\left[\left(X - \mathrm{E}(X)\right)^2\right] = \int_{-\infty}^{\infty} \left(x - \mathrm{E}(X)\right)^2 f(x)\, dx = \int_{-\infty}^{\infty} x^2 f(x) dx - [\mathrm{E}(X)]^2 = E(X^2) - [E(X)]^2.$$

# 1.6.4 Variance of a Continuous Random Variable

**Example 4**

The continuous random variable $X$ is said to have a uniform distribution if its pdf is

$$f(x) = \begin{cases} \frac{1}{b-a} & ; \quad a < x < b \\ 0 & \quad \text{otherwise.} \end{cases}$$




Obtain its mean $\mathrm{E}(X)$ and variance $\mathrm{Var}(X)$.

**Solution**

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} xf(x)dx = \frac{1}{b-a}\int_a^b xdx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b$$

$$= \frac{1}{b-a}\frac{b^2-a^2}{2} = \frac{a+b}{2}.$$

# 1.6.4 Variance of a Continuous Random Variable

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} \left(x - \mathrm{E}(X)\right)^2 f(x)\,dx$$

$$= \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a}\,dx$$

$$= \frac{1}{b-a}\left[\frac{1}{3}\left(x - \frac{a+b}{2}\right)^3\right]_a^b$$

$$= \frac{1}{3(b-a)}\left[\left(b - \frac{a+b}{2}\right)^3 - \left(a - \frac{a+b}{2}\right)^3\right]$$

$$= \frac{1}{3(b-a)}\left[\left(\frac{b-a}{2}\right)^3 - \left(\frac{a-b}{2}\right)^3\right] = \frac{1}{3(b-a)}\left[\left(\frac{b-a}{2}\right)^2\left(\frac{b-a}{2} - \frac{a-b}{2}\right)\right]$$

$$= \frac{(b-a)^2}{12} \quad \blacksquare$$

# 1.6.5 Properties of Mean and Variance

Let $X$ be a random variable and we consider a linear function of $X$, i.e. $Y = aX + b$ with constants $a, b \in \mathbb{R}$. Then

1. $\mathrm{E}(Y) = E(aX + b) = a\mathrm{E}(X) + b$
2. $\mathrm{Var}(Y) = Var(aX + b) = a^2\mathrm{Var}(X)$

# 1.6.5 Properties of Mean and Variance: proof

**Proof** $Y = aX + b$. Then

1) $\mathrm{E}(Y) = \int_{-\infty}^{\infty}(aX + b)f(x)dx = a\int_{-\infty}^{\infty} xf(x)dx + b\int_{-\infty}^{\infty} f(x)dx =$ $aE(x) + b$ (because $\int_{-\infty}^{\infty} f(x)dx = 1$by def.)

2) $\mathrm{Var}(Y) = \int_{-\infty}^{\infty}((aX + b) - E(aX + b))^2\, f(x)dx =$

$$\int_{-\infty}^{\infty}(aX + b - aE(X) - b)^2 f(x)dx$$

$$= \quad a^2\int_{-\infty}^{\infty}\big(X - E(X)\big)^2 f(x)dx = \ a^2\,\mathrm{Var}(X)$$

# 1.6.5 Properties of Mean and Variance: problem

Let $X$ be a random variable with pdf

$$f(x) = \begin{cases} \frac{1}{2}, 0 < x \leq 2 \\ 0 \; elsewhere \end{cases}$$

Show that mean and variance of $Y = 3X \; - 2$ are such that

$E(Y) = 3E(X) - 2$ and $Var(X) = 9Var(X)$

# 1.6.5 Mean of a sum of variables

**Theorem 1**

The mean of a sum of random variables equals the sum of the means, i.e.

$$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$$

This result is always true (for continuous/discrete, and Independent/dependent random variables).

# 1.6.5 Mean of a sum of variables

**Proof** (Continuous case)

We consider $E(X_1 + X_2)$ first.

$$E(X_1 + X_2) = \int_{\mathbb{R}} \int_{\mathbb{R}} (x_1 + x_2) f(x_1, x_2)\, dx_1 dx_2$$

$$= \int_{\mathbb{R}} x_1 \left[\int_{\mathbb{R}} f(x_1, x_2)\, dx_2\right] dx_1 + \int_{\mathbb{R}} x_2 \left[\int_{\mathbb{R}} f(x_1, x_2)\, dx_1\right] dx_2$$

$$= \int_{\mathbb{R}} x_1 f(x_1) dx_1 + \int_{\mathbb{R}} x_2 f(x_2) dx_2 = E(X_1) + E(X_2).$$

This can be further generalized to prove

$$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n).$$

The discrete case can also be similarly proved. ■

# 1.6.5 Mean of a sum of variables, example

- Assume you play a game in which:
  - You flip a coin. If it is head you win 1$, if it is tails you lose 1$
  - You roll a dice: if it is < 4, you lose 1$, if it is 6 you win 2$, otherwise you don't win or lose.
- Find

i) the expected return for one game

ii) The expected return for 5 games

# 1.6.5 Mean of a sum of variables, example solution hard way

- Results and probabilities are

| Coin | | | Dice | | | Tot | | |
|---|---|---|---|---|---|---|---|---|
| Result | Pr | X=W/L | D | Pr | Y=W/L | Pr | Z=W/L | Pr*W/L |
| | | | 1, 2, 3 | 1/2 | -1 | 1/4 | 0 | -1/4 |
| H | 1/2 | 1 | 4. 5 | 1/3 | 0 | 1/6 | 1 | 1/6 |
| | | | 6 | 1/6 | 2 | 1/12 | 3 | 1/4 |
| | | | 1, 2, 3 | 1/2 | -1 | 1/4 | -2 | -1/2 |
| T | 1/2 | -1 | 4. 5 | 1/3 | 0 | 1/6 | -1 | -1/6 |
| | | | 6 | 1/6 | 2 | 1/12 | 1 | 1/12 |

- So the mean is $E(Z) = -\frac{1}{4} + \frac{1}{6} + \frac{1}{4} - \frac{1}{2} - \frac{1}{6} + \frac{1}{12} = -\frac{1}{6}$

# 1.6.5 Mean of a sum of variables, example solution easy way

- Results and probabilities are

| Coin | | | Dice | | | Tot | | |
|---|---|---|---|---|---|---|---|---|
| Result | Pr | X=W/L | D | Pr | Y=W/L | Pr | Z=W/L | Pr*W/L |
| | | | 1, 2, 3 | 1/2 | -1 | 1/4 | 0 | -1/4 |
| H | 1/2 | 1 | 4. 5 | 1/3 | 0 | 1/6 | 1 | 1/6 |
| | | | 6 | 1/6 | 2 | 1/12 | 3 | 1/4 |
| | | | 1, 2, 3 | 1/2 | -1 | 1/4 | -2 | -1/2 |
| T | 1/2 | -1 | 4. 5 | 1/3 | 0 | 1/6 | -1 | -1/6 |
| | | | 6 | 1/6 | 2 | 1/12 | 1 | 1/12 |

- For the coin $E(X) = \frac{1}{2} - \frac{1}{2} = 0,$ for the dice $E(Y) = -\frac{1}{2} + \frac{1}{6} = -\frac{1}{6}$
- $So, E(Z) = E(X) + E(Y) = 0 - \frac{1}{6} = -\frac{1}{6}$

# 1.6.5 Mean of a sum of variables, example solution (ii)

- Results and probabilities are

| Coin | | | Dice | | | Z = X + Y | | |
|---|---|---|---|---|---|---|---|---|
| Result | Pr | X=W/L | D | Pr | Y=W/L | Pr | Z=W/L | Pr*W/L |
| | | | 1, 2, 3 | 1/2 | -1 | 1/4 | 0 | -1/4 |
| H | 1/2 | 1 | 4. 5 | 1/3 | 0 | 1/6 | 1 | 1/6 |
| | | | 6 | 1/6 | 2 | 1/12 | 3 | 1/4 |
| | | | 1, 2, 3 | 1/2 | -1 | 1/4 | -2 | -1/2 |
| T | 1/2 | -1 | 4. 5 | 1/3 | 0 | 1/6 | -1 | -1/6 |
| | | | 6 | 1/6 | 2 | 1/12 | 1 | 1/12 |

- $So, E(Z) = -\frac{1}{6}$. If you play 5 times,

$$E(Z + \cdots + Z) = E(Z) + \cdots + E(Z) = \ 5E(Z) = -5/6$$

# 1.6.6 Summary

- Mean of a random variable
  - $\mu = \mathrm{E}[X] = \sum_{j=1}^{n} x_j \, p(x_j)$ if discrete
  - $\mu = \mathrm{E}[X] = \int_{-\infty}^{\infty} x f(x) dx$ if continuous
- Variance of a random variable
  - $\sigma^2 = \sum_j \left(x_j - \mathrm{E}(X)\right)^2 p(x_j) = \mathrm{E}(\mathrm{X}^2) - \mathrm{E}(X)^2$
  - $\sigma^2 = \displaystyle\int_{-\infty}^{\infty} \left(x - \mathrm{E}(X)\right)^2 f(x)\, dx = \int_{-\infty}^{\infty} x^2 f(x) dx - \mathrm{E}(X)^2$

- $\mathrm{E}(ax + b) = a\mathrm{E}(X) + b$;   $\mathrm{Var}(ax + b) = a^2 \mathrm{Var}(X)$