The Centre and Spread of a Distribution

Sections 7 and 8

Statistics on a Distribution

- Often, important information about a distribution is realized by studying its centre and spread.

- One way to do this is to determine the mean, variance, and standard deviation of the distribution.

Mean or Expected Value

Example: Marks
Consider the following set of marks for 10 students:

Test 1, out of 40:

20, 24, 24, 27, 27, 29, 29, 29, 29, 36

Mark	20	24	27	29	36
Frequency	1	2	2	4	1
Relative Frequency	1/10	2/10	2/10	4/10	1/10

What is the average of these marks?

Mean or Expected Value

Definition:

Let X be a discrete random variable. The mean or the expected value of X is the number

$$E(X) = \sum_x xP(X = x) = \sum_x xp(x)$$

where the sum goes over all values x for which p(x)=P(X=x) is not zero.

Exercise

Example: Leopard Population with Immigration

Consider a population of leopards p_t modelled by

$$p_{t+1} = p_t + I_t \quad \text{where} \quad I_t = \begin{cases} 10 & \text{with a 90\% chance} \\ -100 & \text{with a 10\% chance} \end{cases}$$

Suppose that initially there are 300 leopards. Determine the expected number of leopards after **2** years.

Mean or Expected Value of a Function of a Random Variable

Definition:

Assume that X is a discrete random variable and that p(x)=P(X=x) is its probability mass function. Let g(x) be a function of x. The expected value of the random variable g(X) is

$$E(g(X)) = \sum_{x} g(x)P(X = x) = \sum_{x} g(x)p(x)$$

where the sum goes over all values x for which p(x) is not zero.

Properties of the Expected Value

Theorem:

Let X and Y be discrete random variables and a and b be real numbers. Then

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

(2) $X \pm Y$ is a discrete random variable and

$E(X \pm Y)=E(X) \pm E(Y)$

The Spread of a Distribution

Example: Marks

Consider the following sets of marks for 10 students:

Test 1, out of 40:

20, 20, 20, 20, 20, 40, 40, 40, 40, 40

Test 2, out of 40:

30, 30, 30, 30, 30, 30, 30, 30, 30, 30

Compare the spreads of Test 1 and Test 2 scores.

Variance

Definition:

Assume that X is a random variable with mean $\mu=E(X)$. The variance of X is the real number

$$\sigma^2 = \mathrm{var}(X) = E\left[(X-\mu)^2\right] = E\left[(X-E(X))^2\right]$$

Variance

In words:
The variance of a random variable *X* is the expected value of the difference (squared) between *X* and its mean.

The larger the variance, the larger the spread of a distribution.

Standard Deviation

Definition:

Let *X* be a random variable whose variance is $\sigma^2 = \text{var}(X)$. The standard deviation of *X* is the number

$$\sigma = \sqrt{\text{var}(X)}$$

The standard deviation is measured in the same units as the random variable *X*.

Exercise

Example: Leopard Population with Immigration

Determine the variance and standard deviation for the number of leopards after **2** years.

x	p(x)
100	0.01
210	0.18
320	0.81

Properties of the Variance

Let X be a random variable and a and b be real numbers. Then

(1) $\mathrm{var}(aX+b)=a^2\mathrm{var}(X)$

(2) $\mathrm{var}(X)=E(X^2)-[E(X)]^2$