MH2500 Probability and Introduction to Statistics

Handout 6 - Expected Values - II

Definition

The expected value of a random variable is referred to as a **location parameter** as it indicates the location of the central value of the density or frequency function.

Definition

The expected value of a random variable is referred to as a **location parameter** as it indicates the location of the central value of the density or frequency function.

The median of a distribution is also a location parameter and it is not necessarily equal to the mean.

Definition

The expected value of a random variable is referred to as a **location parameter** as it indicates the location of the central value of the density or frequency function.

The median of a distribution is also a location parameter and it is not necessarily equal to the mean.

The **standard deviation** of a random variable gives an indication of how dispersed the distribution is.

Definition

The expected value of a random variable is referred to as a **location parameter** as it indicates the location of the central value of the density or frequency function.

The median of a distribution is also a location parameter and it is not necessarily equal to the mean.

The **standard deviation** of a random variable gives an indication of how dispersed the distribution is.

Definition

If X is a random variable with expected value $E(X)$, the variance of X is

$$Var(X) = E([X - E(X)]^2)$$

provided that the expectation exist. The standard deviation of X is the square root of the variance.

Variance

If X is a discrete random variable with frequency function $p(X)$ and expected value $\mu = E(X)$, then according to definition and Theorem A of §4.1.1,

$$Var(X) = \sum_i (x_i - \mu)^2 p(x_i).$$

Variance

If X is a discrete random variable with frequency function $p(X)$ and expected value $\mu = E(X)$, then according to definition and Theorem A of §4.1.1,

$$Var(X) = \sum_i (x_i - \mu)^2 p(x_i).$$

If X is a continuous random variable with density function $f(x)$ and $E(X) = \mu$,

$$Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx.$$

Variance

If X is a discrete random variable with frequency function $p(X)$ and expected value $\mu = E(X)$, then according to definition and Theorem A of §4.1.1,

$$Var(X) = \sum_i (x_i - \mu)^2 p(x_i).$$

If X is a continuous random variable with density function $f(x)$ and $E(X) = \mu$,

$$Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx.$$

- Variance is usually denoted σ^2 and standard deviation is denoted σ.

Variance

If X is a discrete random variable with frequency function $p(X)$ and expected value $\mu = E(X)$, then according to definition and Theorem A of §4.1.1,

$$Var(X) = \sum_i (x_i - \mu)^2 p(x_i).$$

If X is a continuous random variable with density function $f(x)$ and $E(X) = \mu$,

$$Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx.$$

- Variance is usually denoted σ^2 and standard deviation is denoted σ.
- $Var(X)$ is the average of $[X - E(X)]^2$, i.e. the square of the deviation of X from its mean.

Variance

If X is a discrete random variable with frequency function $p(X)$ and expected value $\mu = E(X)$, then according to definition and Theorem A of §4.1.1,

$$Var(X) = \sum_i (x_i - \mu)^2 p(x_i).$$

If X is a continuous random variable with density function $f(x)$ and $E(X) = \mu$,

$$Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx.$$

- Variance is usually denoted σ^2 and standard deviation is denoted σ.
- $Var(X)$ is the average of $[X - E(X)]^2$, i.e. the square of the deviation of X from its mean.
- If X has units e.g. meters, then $Var(X)$ has units meter square.

Variance

If X is a discrete random variable with frequency function $p(X)$ and expected value $\mu = E(X)$, then according to definition and Theorem A of §4.1.1,

$$Var(X) = \sum_i (x_i - \mu)^2 p(x_i).$$

If X is a continuous random variable with density function $f(x)$ and $E(X) = \mu$,

$$Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx.$$

- Variance is usually denoted σ^2 and standard deviation is denoted σ.
- $Var(X)$ is the average of $[X - E(X)]^2$, i.e. the square of the deviation of X from its mean.
- If X has units e.g. meters, then $Var(X)$ has units meter square.
- We are ultimately interested in the standard deviation, but usually we compute the variance first then take square root.

Variance under linear transformation

Theorem A

If $Var(X)$ exists and $Y = a + bX$, then $Var(Y) = b^2 Var(X)$.

. . . .

Variance under linear transformation

Theorem A

If $Var(X)$ exists and $Y = a + bX$, then $Var(Y) = b^2 Var(X)$.

Proof:

Clearly $E(Y) = a + bE(X)$.

Variance under linear transformation

Theorem A

If $Var(X)$ exists and $Y = a + bX$, then $Var(Y) = b^2 Var(X)$.

Proof:

Clearly $E(Y) = a + bE(X)$. Therefore,

$$
\begin{aligned}
E([Y - E(Y)]^2) &= E([a + bX - (a + bE(X))]^2) \\
&= \\
&= \\
&= \qquad\qquad .
\end{aligned}
$$

Theorem A

If $Var(X)$ exists and $Y = a + bX$, then $Var(Y) = b^2 Var(X)$.

Proof:

Clearly $E(Y) = a + bE(X)$. Therefore,

$$\begin{aligned} E([Y - E(Y)]^2) &= E([a + bX - (a + bE(X))]^2) \\ &= E([a + bX - a - bE(X)]^2) \\ &= \\ &= \qquad . \end{aligned}$$

Variance under linear transformation

Theorem A

If $Var(X)$ exists and $Y = a + bX$, then $Var(Y) = b^2 Var(X)$.

Proof:

Clearly $E(Y) = a + bE(X)$. Therefore,

$$\begin{aligned} E([Y - E(Y)]^2) &= E([a + bX - (a + bE(X))]^2) \\ &= E([a + bX - a - bE(X)]^2) \\ &= E(b^2(X - E(X)]^2) \\ &= \qquad\qquad . \end{aligned}$$

Variance under linear transformation

Theorem A

If $Var(X)$ exists and $Y = a + bX$, then $Var(Y) = b^2 Var(X)$.

Proof:

Clearly $E(Y) = a + bE(X)$. Therefore,

$$\begin{aligned} E([Y - E(Y)]^2) &= E([a + bX - (a + bE(X))]^2) \\ &= E([a + bX - a - bE(X)]^2) \\ &= E(b^2(X - E(X)]^2) \\ &= b^2 Var(X). \end{aligned}$$

Variance under linear transformation

Theorem A

If $Var(X)$ exists and $Y = a + bX$, then $Var(Y) = b^2 Var(X)$.

Proof:

Clearly $E(Y) = a + bE(X)$. Therefore,

$$\begin{aligned} E([Y - E(Y)]^2) &= E([a + bX - (a + bE(X))]^2) \\ &= E([a + bX - a - bE(X)]^2) \\ &= E(b^2(X - E(X)]^2) \\ &= b^2 Var(X). \end{aligned}$$

Remarks: Adding a constant does not change variance since it measures spread around a center. In symbols, it means $\sigma_Y = |b|\sigma_X$.

Example A

Bernoulli Distribution

Suppose X has a Bernoulli distribution, i.e., X takes values 0 and 1 with probability $1 - p$ and p, respectively. Find the variance of X. For which values of p is the variance maximum? minimum?

. . . .

Example A

Bernoulli Distribution

Suppose X has a Bernoulli distribution, i.e., X takes values 0 and 1 with probability $1-p$ and p, respectively. Find the variance of X. For which values of p is the variance maximum? minimum?

. . . .

From Section 4.1.2, $E(X) = p$. By definition of variance,

$$
\begin{aligned}
Var(X) &= \\
&= \\
&= \qquad\qquad .
\end{aligned}
$$

Example A

Bernoulli Distribution

Suppose X has a Bernoulli distribution, i.e., X takes values 0 and 1 with probability $1-p$ and p, respectively. Find the variance of X. For which values of p is the variance maximum? minimum?

. . . .

From Section 4.1.2, $E(X) = p$. By definition of variance,

$$\begin{aligned} Var(X) &= (0-p)^2 \times (1-p) + (1-p)^2 \times p \\ &= \\ &= \qquad . \end{aligned}$$

Example A

Bernoulli Distribution

Suppose X has a Bernoulli distribution, i.e., X takes values 0 and 1 with probability $1-p$ and p, respectively. Find the variance of X. For which values of p is the variance maximum? minimum?

. . . .

From Section 4.1.2, $E(X) = p$. By definition of variance,

$$\begin{aligned} Var(X) &= (0-p)^2 \times (1-p) + (1-p)^2 \times p \\ &= p^2(1-p) + (1-p)^2 p \\ &= \qquad\qquad . \end{aligned}$$

Example A

Bernoulli Distribution

Suppose X has a Bernoulli distribution, i.e., X takes values 0 and 1 with probability $1 - p$ and p, respectively. Find the variance of X. For which values of p is the variance maximum? minimum?

. . . .

From Section 4.1.2, $E(X) = p$. By definition of variance,

$$\begin{aligned} Var(X) &= (0-p)^2 \times (1-p) + (1-p)^2 \times p \\ &= p^2(1-p) + (1-p)^2 p \\ &= p(1-p). \end{aligned}$$

Example A

Bernoulli Distribution

Suppose X has a Bernoulli distribution, i.e., X takes values 0 and 1 with probability $1-p$ and p, respectively. Find the variance of X. For which values of p is the variance maximum? minimum?

From Section 4.1.2, $E(X) = p$. By definition of variance,

$$\begin{aligned} Var(X) &= (0-p)^2 \times (1-p) + (1-p)^2 \times p \\ &= p^2(1-p) + (1-p)^2 p \\ &= p(1-p). \end{aligned}$$

When $p = \frac{1}{2}$, the expression $p(1-p)$ attains its maximum.

Example A

Bernoulli Distribution

Suppose X has a Bernoulli distribution, i.e., X takes values 0 and 1 with probability $1-p$ and p, respectively. Find the variance of X. For which values of p is the variance maximum? minimum?

From Section 4.1.2, $E(X) = p$. By definition of variance,

$$\begin{aligned} Var(X) &= (0-p)^2 \times (1-p) + (1-p)^2 \times p \\ &= p^2(1-p) + (1-p)^2 p \\ &= p(1-p). \end{aligned}$$

When $p = \frac{1}{2}$, the expression $p(1-p)$ attains its maximum.
When $p = 0$ or 1, the expression $p(1-p) = 0$ and variance is at its minimum.

Example B -Normal distribution

Suppose X has density function

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

We have previously shown that $E(X) = \mu$. Now, show that $Var(X) = \sigma^2$.

. . . .

Example B -Normal distribution

Suppose X has density function

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

We have previously shown that $E(X) = \mu$. Now, show that $Var(X) = \sigma^2$.

. . . .

$$Var(X) = E[(X-\mu)^2] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

Example B -Normal distribution

Suppose X has density function

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

We have previously shown that $E(X) = \mu$. Now, show that $Var(X) = \sigma^2$.

. . . .

$$Var(X) = E[(X-\mu)^2] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

Making the change of variable $z = (x-\mu)/\sigma$, the right hand side becomes

$$\frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-z^2/2} dz.$$

Example B -Normal distribution

Suppose X has density function

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

We have previously shown that $E(X) = \mu$. Now, show that $Var(X) = \sigma^2$.

. . . .

$$Var(X) = E[(X-\mu)^2] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

Making the change of variable $z = (x-\mu)/\sigma$, the right hand side becomes

$$\frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-z^2/2} dz.$$

Making a further change of variable $u = z^2/2$ reduces the integral to a gamma function.
Hence $Var(X) = \sigma^2$.

Theorem

The variance of X, if it exists, may also be calculated as follows:

$$Var(X) = E(X^2) - [E(X)]^2.$$

Theorem

The variance of X, if it exists, may also be calculated as follows:

$$Var(X) = E(X^2) - [E(X)]^2.$$

Proof:

Let μ denote $E(X)$. Then

$$Var(X) =$$

Theorem

The variance of X, if it exists, may also be calculated as follows:

$$Var(X) = E(X^2) - [E(X)]^2.$$

Proof:

Let μ denote $E(X)$. Then

$$Var(X) = E[(X - \mu)^2] =$$

Theorem

The variance of X, if it exists, may also be calculated as follows:

$$Var(X) = E(X^2) - [E(X)]^2.$$

Proof:

Let μ denote $E(X)$. Then

$$Var(X) = E[(X-\mu)^2] = E(X^2 - 2X\mu + \mu^2).$$

Theorem

The variance of X, if it exists, may also be calculated as follows:

$$Var(X) = E(X^2) - [E(X)]^2.$$

Proof:

Let μ denote $E(X)$. Then

$$Var(X) = E[(X-\mu)^2] = E(X^2 - 2X\mu + \mu^2).$$

Applying Theorem A (linearity),

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

Theorem

The variance of X, if it exists, may also be calculated as follows:

$$Var(X) = E(X^2) - [E(X)]^2.$$

Proof:

Let μ denote $E(X)$. Then

$$Var(X) = E[(X-\mu)^2] = E(X^2 - 2X\mu + \mu^2).$$

Applying Theorem A (linearity),

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

Theorem

The variance of X, if it exists, may also be calculated as follows:

$$Var(X) = E(X^2) - [E(X)]^2.$$

Proof:
Let μ denote $E(X)$. Then

$$Var(X) = E[(X-\mu)^2] = E(X^2 - 2X\mu + \mu^2).$$

Applying Theorem A (linearity),

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

Example C

Suppose X is uniform on [0,1]. Find $Var(X)$.

. . . .

Example C

Suppose X is uniform on [0,1]. Find $Var(X)$.

. . . .

Clearly $E(X) = \frac{1}{2}$.

$$E(X^2) = \qquad\qquad\qquad .$$

Therefore,

$$Var(X) = \qquad\qquad\qquad .$$

Example C

Suppose X is uniform on [0,1]. Find $Var(X)$.

. . . .

Clearly $E(X) = \frac{1}{2}$.

$$E(X^2) = \int_0^1 x^2 dx = \frac{1}{3}.$$

Therefore,

$$Var(X) = \qquad\qquad\qquad .$$

Example C

Suppose X is uniform on [0,1]. Find $Var(X)$.

. . . .

Clearly $E(X) = \frac{1}{2}$.

$$E(X^2) = \int_0^1 x^2 dx = \frac{1}{3}.$$

Therefore,

$$Var(X) = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}.$$

Chebyshev's inequality

Let X be a random variable with mean μ and variance σ^2. Then for any $t > 0$,

$$P(|X - \mu| > t) \leq \frac{\sigma^2}{t^2}.$$

Chebyshev's inequality

Let X be a random variable with mean μ and variance σ^2. Then for any $t > 0$,

$$P(|X - \mu| > t) \leq \frac{\sigma^2}{t^2}.$$

Proof: Let $Y = (X - \mu)^2$. Then $E(Y) = Var(X) = \sigma^2$.

Chebyshev's inequality

Let X be a random variable with mean μ and variance σ^2. Then for any $t > 0$,

$$P(|X - \mu| > t) \leq \frac{\sigma^2}{t^2}.$$

Proof: Let $Y = (X - \mu)^2$. Then $E(Y) = Var(X) = \sigma^2$.

Recall Markov's inequality

If X is a random variable with $P(X \geq 0) = 1$ and for which $E(X)$ exists, then $P(X \geq t) \leq E(X)/t$.

Chebyshev's inequality

Let X be a random variable with mean μ and variance σ^2. Then for any $t > 0$,

$$P(|X - \mu| > t) \leq \frac{\sigma^2}{t^2}.$$

Proof: Let $Y = (X - \mu)^2$. Then $E(Y) = Var(X) = \sigma^2$.

Recall Markov's inequality

$\boxed{\text{If } X \text{ is a random variable with } P(X \geq 0) = 1 \text{ and for which } E(X) \text{ exists, then } P(X \geq t) \leq E(X)/t.}$

Apply it to Y and t^2. We get

$$= P(Y \geq t^2) \leq \frac{E(Y)}{t^2} = \quad .$$

Chebyshev's inequality

Let X be a random variable with mean μ and variance σ^2. Then for any $t > 0$,

$$P(|X - \mu| > t) \leq \frac{\sigma^2}{t^2}.$$

Proof: Let $Y = (X - \mu)^2$. Then $E(Y) = Var(X) = \sigma^2$.

Recall Markov's inequality

> If X is a random variable with $P(X \geq 0) = 1$ and for which $E(X)$ exists, then $P(X \geq t) \leq E(X)/t$.

Apply it to Y and t^2. We get

$$P(|X - \mu| > t) \;=\; P(Y \geq t^2) \;\leq\; \frac{E(Y)}{t^2} \;=\; \quad .$$

Chebyshev's inequality

Let X be a random variable with mean μ and variance σ^2. Then for any $t > 0$,

$$P(|X - \mu| > t) \leq \frac{\sigma^2}{t^2}.$$

Proof: Let $Y = (X - \mu)^2$. Then $E(Y) = Var(X) = \sigma^2$.

Recall Markov's inequality

If X is a random variable with $P(X \geq 0) = 1$ and for which $E(X)$ exists, then $P(X \geq t) \leq E(X)/t$.

Apply it to Y and t^2. We get

$$P(|X - \mu| > t) \;=\; P(Y \geq t^2) \;\leq\; \frac{E(Y)}{t^2} \;=\; \frac{\sigma^2}{t^2}.$$

Consequences of Chebyshev's inequality

- If σ^2 is very small, then the probability of X deviating from μ by t is very small. Hence the probability that X deviates from μ by not much is very high.

Consequences of Chebyshev's inequality

- If σ^2 is very small, then the probability of X deviating from μ by t is very small. Hence the probability that X deviates from μ by not much is very high.

- Set $t = k\sigma$ so that we get
$$P(|X - \mu| > k\sigma) \leq \frac{1}{k^2}.$$
Thus the probability that X is more than $k\sigma$ away from μ is less than $\frac{1}{k^2}$. E.g. $k = 2$, then $P(|x - \mu| > 2\sigma) \leq \frac{1}{4}$. For a particular distribution, the probability could be lower. E.g. For a normal distribution, $P(|X - \mu| > 2\sigma) \approx 1/20$.

Consequences of Chebyshev's inequality

- If σ^2 is very small, then the probability of X deviating from μ by t is very small. Hence the probability that X deviates from μ by not much is very high.

- Set $t = k\sigma$ so that we get
$$P(|X - \mu| > k\sigma) \leq \frac{1}{k^2}.$$
Thus the probability that X is more than $k\sigma$ away from μ is less than $\frac{1}{k^2}$. E.g. $k = 2$, then $P(|x - \mu| > 2\sigma) \leq \frac{1}{4}$. For a particular distribution, the probability could be lower. E.g. For a normal distribution, $P(|X - \mu| > 2\sigma) \approx 1/20$.

- These results hold for any random variable with distribution provided the variance exists.

Corollary A

If $Var(X) = 0$, then $P(X = \mu) = 1$.

Corollary A

If $Var(X) = 0$, then $P(X = \mu) = 1$.

Proof:
Suppose not, that is, suppose $Var(X) = 0$ and $P(X = \mu) < 1$.

Corollary A

If $Var(X) = 0$, then $P(X = \mu) = 1$.

Proof:

Suppose not, that is, suppose $Var(X) = 0$ and $P(X = \mu) < 1$. Therefore, for some $\epsilon > 0$,

$$P(|X - \mu| \geq \epsilon) > 0.$$

Corollary A

If $Var(X) = 0$, then $P(X = \mu) = 1$.

Proof:
Suppose not, that is, suppose $Var(X) = 0$ and $P(X = \mu) < 1$.
Therefore, for some $\epsilon > 0$,

$$P(|X - \mu| \geq \epsilon) > 0.$$

This contradicts Chebyshev's inequality which implies that for any $\epsilon > 0$,

$$P(|X - \mu| \geq \epsilon) \leq \frac{Var(X)}{\epsilon^2} = 0.$$

Example D – continuing from E.g in Handout 5 slide 25/25

An investor wishes to apportion a capital on two investments, one is risky (e.g. stocks) and other is risk free (e.g. savings account). The investments have returns R_1 and R_2, which are modelled as random variables with expectations μ_1, and μ_2 where $\mu_1 > \mu_2$, and standard deviations σ_1 and σ_2.

What is σ_2?

The investor apportions a fraction π of her capital on the first investment and $(1 - \pi)$ of her capital on the second risk free investment. Recall that the returns is $R = \pi R_1 + (1 - \pi)R_2$. What is $E(R)$ and $Var(R)$?

How should an investor choose π?

Example D con't

Since the second investment is risk free, $Var(R_2) = 0$.

$$E(R) = \pi\mu_1 + (1 - \pi)\mu_2.$$

Example D con't

Since the second investment is risk free, $Var(R_2) = 0$.

$$E(R) = \pi\mu_1 + (1-\pi)\mu_2.$$

Since $Var(R_2) = 0$, we treat R_2 like a constant.

$$Var(R) =$$

Example D con't

Since the second investment is risk free, $Var(R_2) = 0$.

$$E(R) = \pi\mu_1 + (1-\pi)\mu_2.$$

Since $Var(R_2) = 0$, we treat R_2 like a constant.

$$Var(R) = \pi^2 Var(R_1) = \pi^2\sigma_1^2.$$

Example D con't

Since the second investment is risk free, $Var(R_2) = 0$.

$$E(R) = \pi\mu_1 + (1-\pi)\mu_2.$$

Since $Var(R_2) = 0$, we treat R_2 like a constant.

$$Var(R) = \pi^2 Var(R_1) = \pi^2\sigma_1^2.$$

The larger π is, the larger the expected return but the larger risk. An investor has to balance the risk she is willing to take against the expected gain. The desired balance is different for different investor.