Variance

Lecture 5a (STAT 24400 F24)

Variance of random variables

The **variance** of a random variable X is defined as

$$\mathrm{Var}(X) = \mathbb{E}\Big((X - \mathbb{E}(X))^2\Big)$$

- Variance measures how much the r.v. typically varies from its mean.
- The definition is valid if and only if the expectation exists, and $\mathbb{E}(X)$ exists.
- Variance is often denoted as σ^2 or σ_X^2.
- σ or σ_X is called the **standard deviation** of X.

Variance of discrete and continuous r.v.'s

For a discrete random variable, we may calculate the variance as

$$\mathrm{Var}(X) = \sum_x p_X(x) \cdot (x - \mu_X)^2$$

For a continuous random variable,

$$\mathrm{Var}(X) = \int_{x=-\infty}^{\infty} f_X(x) \cdot (x - \mu_X)^2 \, \mathrm{d}x$$

Variance and measures of variation

Why use $\mathbb{E}\Big((X - \mathbb{E}(X))^2\Big)$ rather than, say, $\mathbb{E}\Big(\big|X - \mathbb{E}(X)\big|\Big)$?

- Intuitively, both measure the same thing:
  how much variability is in the distribution of X?
- However, $\mathbb{E}\Big((X - \mathbb{E}(X))^2\Big)$ has many convenient mathematical and statistical properties (because mathematically x^2 is a nicer function than $|x|$).

1. **A simpler formula.**
   Variance can be calculated with a simpler formula:

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

(Note: this also shows that $\mathbb{E}(g(X)) \neq g(\mathbb{E}(X))$ in general.)

Proof:. Write $\mu_X = \mathbb{E}(X)$,

$$\begin{aligned}\mathrm{Var}(X) = \mathbb{E}((X-\mu_X)^2) &= \mathbb{E}(X^2 - 2\mu_X \cdot X + \mu_X^2)\\ &= \mathbb{E}(X^2) - 2\mu_X\mathbb{E}(X) + \mu_X^2 = \mathbb{E}(X^2) - \mu_X^2\end{aligned}$$

by linearity of $\mathbb{E}(\cdot)$

Properties of variance

2. **The degenerate case.**

$\mathrm{Var}(X) = 0$ if and only if $\mathbb{P}(X = \mu_X) = 1$.

Properties of variance

3. **Chebyshev's inequality.**

For any random variable X and any $t > 0$,

$$\mathbb{P}(|X - \mu_X| \geq t) \leq \frac{\sigma_X^2}{t^2}$$

Proof.

Recall Markov's inequality, for $X^* \geq 0$, $\quad \mathbb{P}(X^* \geq t) \leq \frac{\mu_{X^*}}{t}$.

Let $Y = (X - \mu_X)^2$, so then $\mathbb{E}(Y) = \sigma_X^2$.

$$\mathbb{P}(|X - \mu_X| \geq t) = \mathbb{P}(Y \geq t^2) \leq \frac{\sigma_X^2}{t^2}$$

by Markov's inequality applied to Y

Properties of variance

4. **Linear transformations.**

For any random variable X and any constants a, b,

$$\mathrm{Var}(a + bX) = b^2\,\mathrm{Var}(X)$$

Proof.

Using linearity of $\mathbb{E}(\cdot)$,

$$\begin{aligned}\mathrm{Var}(a+bX) &= \mathbb{E}[(a + bX - \mu_{a+bX})^2]\\ &= \mathbb{E}[(a + bX - (a + b\mu_X))^2]\\ &= \mathbb{E}[(b(X - \mu_X))^2] = b^2\mathbb{E}[(X-\mu_X)^2] = b^2\,\mathrm{Var}(X)\end{aligned}$$

Examples (Normal, by formula)

Suppose $Y \sim N(\mu, \sigma^2)$. Formally, we may use the formulae to calculate

$$\mathbb{E}(Y) = \int_{-\infty}^{\infty} y\, f_Y(y) dy, \quad \mathrm{Var}(Y) = \mathbb{E}(Y - \mu_Y)^2) = \int_{-\infty}^{\infty} (y-\mu)^2 f_Y(y) dy$$

Applying variable substitution $z = \frac{y-\mu}{\sigma}$ and Integration-by-parts,

$$\mathbb{E}(Y) = \int_{-\infty}^{\infty} y \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy = \quad \cdots \quad = \mu$$

$$\mathrm{Var}(Y) = \int_{-\infty}^{\infty} (y-\mu)^2 \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy = \quad \cdots \quad = \sigma^2$$

(Side notes: Properties of odd and even functions.)

Examples (Normal, by linearity)

Alternatively:

The standard normal distribution $\mathrm{N}(0,1)$ has mean 0 and variance 1 (assumption).

Let $X \sim \mathrm{N}(0,1)$. Then $Y = \mu + \sigma \cdot X$ is also of normal distribution (assumption).

Verify: $Y \sim \mathrm{N}(\mu, \sigma^2)$.

Using the linearity property of expected value $\mathbb{E}(\cdot)$:

$$\mathbb{E}(Y) = \mathbb{E}(\mu + \sigma \cdot X) = \mu + \sigma \cdot \mathbb{E}(X) = \mu$$

$$\mathrm{Var}(Y) = \mathrm{Var}(\mu + \sigma \cdot X) = \sigma^2 \cdot \mathrm{Var}(X) = \sigma^2$$

Examples (Bernoulli)

Suppose $X \sim$Bernoulli(p). What is $\mathrm{Var}(X)$?

We know that

$$\mathbb{E}(X) = p$$

Compute

$$\mathbb{E}(X^2) = 1^2 \cdot p + 0^2 \cdot (1-p) = p$$

So,

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = p - p^2 = p(1-p)$$

Remark: Note the useful symmetry in the variance.

Examples (Exponential)

Suppose $X \sim$Exponential(λ). What is $\mathrm{Var}(X)$?

We know that

$$\mathbb{E}(X) = 1/\lambda$$

Applying Integration-by-parts twice,

$$\mathbb{E}(X^2) = \int_{x=0}^{\infty} x^2 \cdot \lambda e^{-\lambda x}\, \mathrm{d}x = \left[-x^2 e^{-\lambda x} - \frac{2}{\lambda} x e^{-\lambda x} - \frac{2}{\lambda^2} e^{-\lambda x} \right]_{x=0}^{\infty} = \frac{2}{\lambda^2}$$

So,

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

Example: sampling effects

We perform and compare two experiments:

- Toss a fair coin 2 times, and let $X = \#$ of Heads
- A bag has 4 tickets (2 red, 2 blue). Draw 2 tickets, and let $Y = \#$ red

<u>Questions</u>

Are $\mathbb{E}(X)$ and $\mathbb{E}(Y)$ the same?

Are $\text{Var}(X)$ and $\text{Var}(Y)$ the same?

Examples (sampling effects, cont.)

Calculate the probability mass functions of X and Y:

	0	1	2	
p_X	1/4	1/2	1/4	⟵ using Binomial PMF
p_Y	1/6	2/3	1/6	⟵ by counting

$$\mathbb{E}(X) = \tfrac{1}{4} \cdot 0 + \tfrac{1}{2} \cdot 1 + \tfrac{1}{4} \cdot 2 = 1$$
$$\mathbb{E}(Y) = \tfrac{1}{6} \cdot 0 + \tfrac{2}{3} \cdot 1 + \tfrac{1}{6} \cdot 2 = 1$$

$$\mathbb{E}(X^2) = \tfrac{1}{4} \cdot 0^2 + \tfrac{1}{2} \cdot 1^2 + \tfrac{1}{4} \cdot 2^2 = \tfrac{3}{2} \quad \rightsquigarrow \quad \text{Var}(X) = \tfrac{3}{2} - 1^2 = \tfrac{1}{2}$$
$$\mathbb{E}(Y^2) = \tfrac{1}{6} \cdot 0^2 + \tfrac{2}{3} \cdot 1^2 + \tfrac{1}{6} \cdot 2^2 = \tfrac{4}{3} \quad \rightsquigarrow \quad \text{Var}(Y) = \tfrac{4}{3} - 1^2 = \tfrac{1}{3}$$

Examples (remarks)

<u>Remarks</u>

Notice the differences in the two experiments:

- Independent vs not-independent sampling
- without replace vs. with replacement
- infinite population vs finite population

$\Rightarrow$ The effects on variance