Last topic: Expected value of a RV

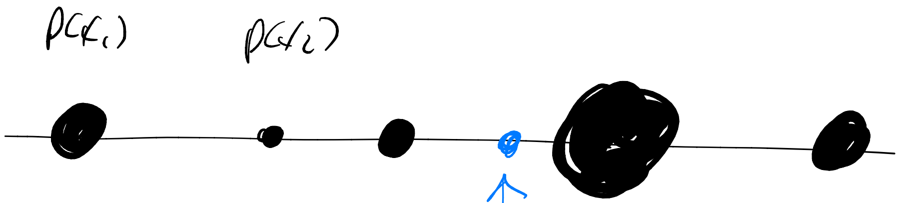
Idea: longterm average

X discr. $\quad \sum_i x_i \, p(x_i) \quad , \; x_i$: poss. values of X

X cont. $\quad \int_{\mathbb{R}} x \cdot f(x) \, dx \quad , \; f$ pdf

Another interpretation: center of gravity

X discr.: a number of points $x_1, x_2, \ldots$, with weight $p(x_i)$

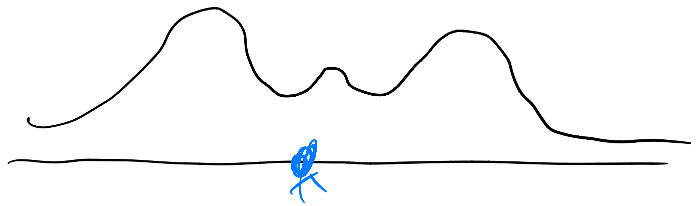


points on a weightless rod

center of gravity supports the construction

X cont

weight proportional to weight of curve



center of gravity

Properties of Exp. V.:

$$E[a\mathcal{X}] = a\,E[\mathcal{X}]$$

$$E[b] = b$$

$$E[\mathcal{X} + \mathcal{Y}] = E[\mathcal{X}] + E[\mathcal{Y}]$$

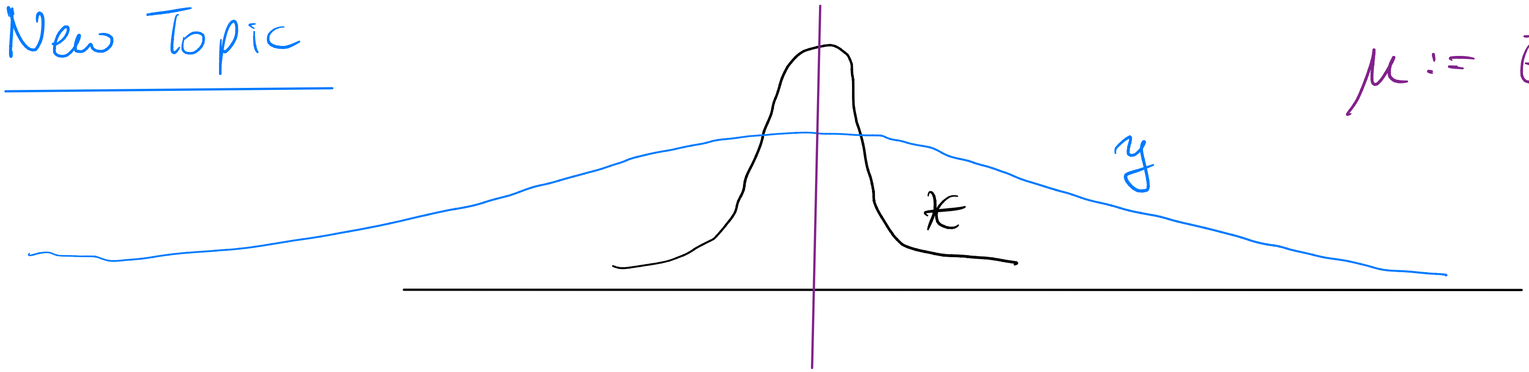
! holds also if $\mathcal{X}, \mathcal{Y}$ are <u>not</u> independent

$$E[\mathcal{X}_1 + \dots + \mathcal{X}_n] = \sum_{i=1}^{n} E[\mathcal{X}_i]$$

LOTUS

$$E[g(\mathcal{X})] = \int_{\mathbb{R}} g(x)\, f(x)\, dx$$

New Topic



$\mu := E[X] = E[Y]$

μ mean of X, Y

Values of Y are spread much farther around μ than the values of X.

2.6 Variance

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

$$g(x) = (x-\mu)^2$$

is the variance of X.

Why not $E[|X-\mu|]$?

The definition with the square has better mathematical properties.

How can we calculate $Var(X)$? Suppose $X \sim f$.

1.) Apply LOTUS:

$$Var(X) = \int_{\mathbb{R}} (x-\mu)^2 f(x)\,dx$$

$$= \int_{\mathbb{R}} (x^2 - 2x\mu + \mu^2) f(x)\,dx$$

$$= \int_{\mathbb{R}} x^2 f(x)\,dx - 2\mu \int_{\mathbb{R}} x f(x)\,dx + \mu^2 \int_{\mathbb{R}} f(x)\,dx$$

$\boxed{\text{Shortcut}}$

$$= E[X^2] - 2\mu E[X] + \mu^2 \cdot 1$$

$$= E[X^2] - 2\mu^2 + \mu^2$$

$$= E[X^2] - E[X]^2$$

square of mean

2nd moment of X

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

Example 44 $\quad X$ = number on die

$E[X] = \frac{7}{2}$. $\quad$ Also $\quad E[X^2] = \frac{91}{6}$ $\quad$ (Ex. 37)

$$Var(X) = E[X^2] - E[X]^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{182 - 3 \cdot 7^2}{12}$$
$$= \frac{182 - 147}{12} = \frac{35}{12}$$

Units: unit of measurement of X is metre, sec

$\Rightarrow$ unit of measurement of X^2 is metre^2, sec^2

Get back to original unit: take $\sqrt{\cdot}$:

$$\sigma := \sqrt{Var(X)}$$

is the standard deviation.

One often writes the variance as the square of the standard deviation:

$$Var(X) = \sigma^2$$

Properties of $\mathrm{Var}(\cdot)$.

Suppose X has $E[X] = \mu$.

- $\mathrm{Var}(X + b) = \mathrm{Var}(X)$
- $\mathrm{Var}(aX) = a^2 \, \mathrm{Var}(X)$
- $\mathrm{Var}(X + Y) \overset{?}{=} \mathrm{Var}(X) + \mathrm{Var}(Y)$ only if X, Y independent

$$\mathrm{Var}(\overbrace{aX + b}^{Y}) = E[Y^2] - E[Y]^2$$

$$= E[(aX+b)^2] - E[aX+b]^2$$

$$= E[a^2X^2 + 2abX + b^2] - (aE[X] + b)^2$$

$$= a^2 E[X^2] + 2ab E[X] + E[b^2]$$

$$- a^2\mu^2 - 2ab\mu - b^2$$

$$= a^2 E[X^2] + 2ab E[X] + E[b^2] - a^2\mu^2 - 2ab\mu - b^2$$

$$= a^2 E[X^2] + \cancel{2ab\mu} + \cancel{b^2} - a^2\mu^2 - \cancel{2ab\mu} - \cancel{b^2}$$

$$= a^2\left(E[X^2] - \mu^2\right) = a^2 \, Var(X)$$

Note: $\sigma_{aX} = a\,\sigma_X$

2.7 Covariance

We note

$$Var(X+X) = Var(2X) = 4\,Var(X)$$
$$\neq Var(X) + Var(X)$$

<u>Definition 45:</u> X, Y RV, $\mu_X = E[X]$, $\mu_Y = E[Y]$

Then

$$Cov(X,Y) = E\left[(X-\mu_X)(Y-\mu_Y)\right]$$

This assumes a joint distribution of X and Y.

Property of Cov

$$\mathrm{Cov}(X,Y) = E\left[(X-\mu_X)(Y-\mu_Y)\right]$$

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= E\left[XY - \mu_X Y - \mu_Y X + \mu_X\mu_Y\right] \\
&= E[XY] - \mu_X E[Y] - \mu_Y E[X] + \mu_X \mu_Y \\
&= E[XY] - \mu_X\mu_Y - \cancel{\mu_Y\mu_X} + \cancel{\mu_X\mu_Y} \\
&= E[XY] - E[X]E[Y]
\end{aligned}$$

Observation: X, Y independent $\Rightarrow$ $E[XY] = E[X]\cdot E[Y]$

$\Rightarrow$ $\mathrm{Cov}(X,Y) = 0$

Proposition 46 $\quad Cov(X_1+X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y)$

$$Cov(X_1+X_2, Y) = E[(X_1+X_2)\cdot Y] - E[X_1+X_2]\cdot E[Y]$$

$$= E[X_1 Y] + E[X_2 Y] - E[X_1]\cdot E[Y] - E[X_2]\cdot E[Y]$$

$$= E[X_1 Y] - E[X_1]\cdot E[Y]$$

$$+ E[X_2 Y] - E[X_2]\cdot E[Y]$$

$$= Cov(X_1, Y) + Cov(X_2, Y)$$

Theorem 47

$$Cov\left(\sum_i X_i, \sum_j Y_j\right) = \sum_i \sum_j Cov(X_i, Y_j)$$

Consequence:

$$\mathrm{Var}\Big(\sum_i X_i\Big) = \mathrm{Cov}\Big(\sum_i X_i, \sum_i X_i\Big)$$

$$= \sum_i \sum_j \mathrm{Cov}(X_i, X_j)$$

$$= \sum_i \Big(\sum_{j \neq i} \mathrm{Cov}(X_i, X_j) + \mathrm{Cov}(X_i, X_i) \Big)$$

$$= \Big(\sum_i \sum_{j \neq i} \mathrm{Cov}(X_i, X_j) \Big) + \sum_i \mathrm{Cov}(X_i, X_i)$$

$$= \sum_i \mathrm{Var}(X_i) + \sum_i \sum_{j \neq i} \mathrm{Cov}(X_i, X_j)$$

For $n=2$:

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X, Y)$$

Observation: X, Y ind. $\Rightarrow E[XY] = E[X] \cdot E[Y]$

Suppose $X \sim f$, $Y \sim g$:

LOTUS

$$E[XY] = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} x \cdot y \, f(x) g(y) \, dx \right) dy$$

$$= \int_{\mathbb{R}} y \cdot g(y) \left(\int_{\mathbb{R}} x \cdot f(x) \, dx \right) dy$$

$$= \int_{\mathbb{R}} x \cdot f(x) \, dx \cdot \int_{\mathbb{R}} y \cdot g(y) \, dy$$

$$= E[X] \cdot E[Y]$$

Meaning of covariance $Cov(X, Y)$

> 0: X, Y grow together in the same direction

< 0: X, Y grow in sync in opposite directions

≈ 0: X, Y vary independently

Normalize RVs X, Y by taking $\frac{X}{\sigma_X}, \frac{Y}{\sigma_Y}$,

i.e., $\frac{X}{\sqrt{Var(X)}} \quad \frac{Y}{\sqrt{Var(Y)}}$

$$\boxed{Corr(X, Y)} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = \frac{Cov(X, Y)}{\sqrt{Var(X) \cdot Var(Y)}}$$

Correlation between X and Y

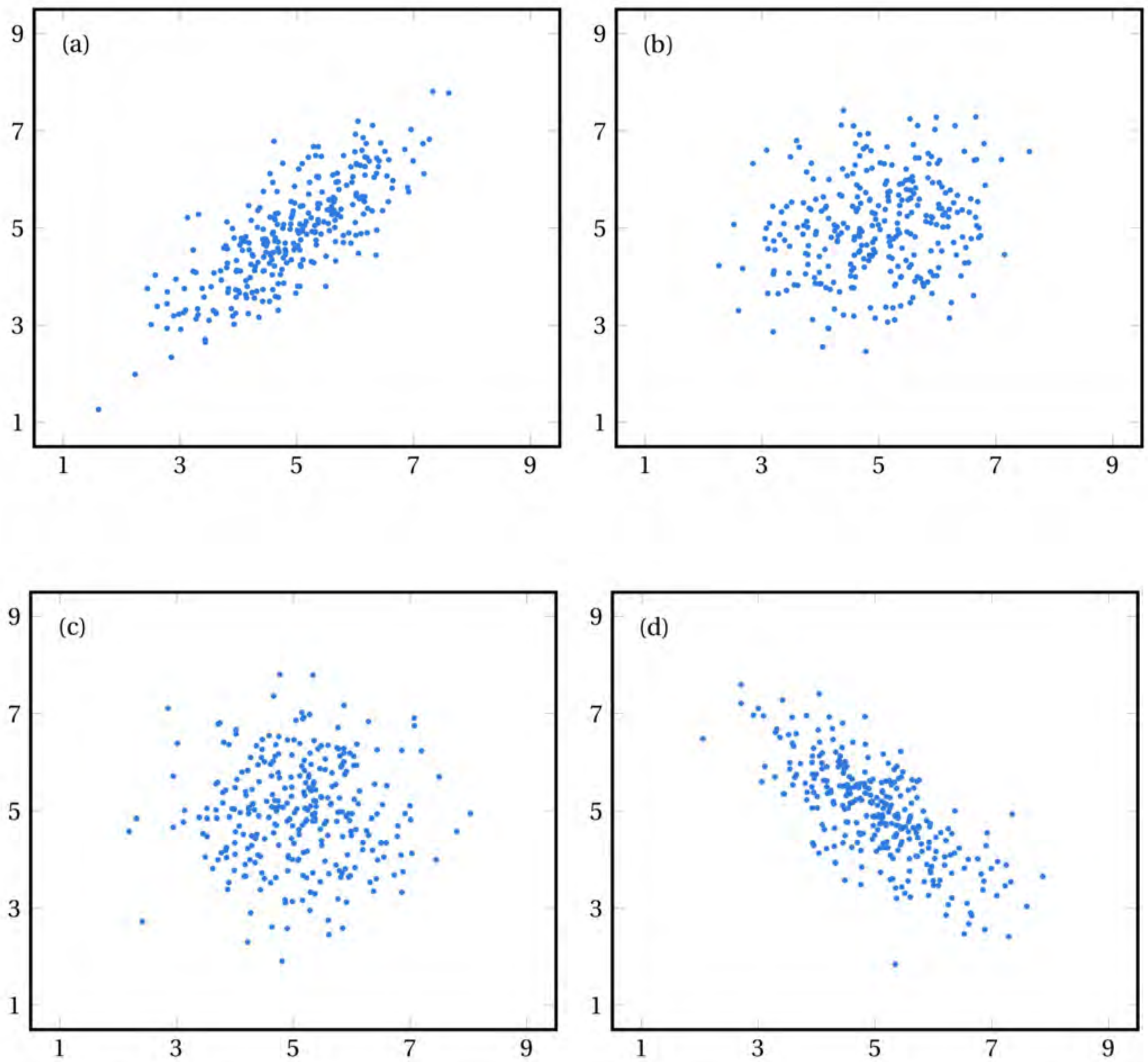


Figure 9: Random variables $\mathscr{X}$ and $\mathscr{Y}$ with correlations (a) 0.75; (b) 0.2; (c) 0; and (d) −0.75.

Example: 10 independent dice rolls: X_i is i-th roll

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) \underset{\text{ind}}{=} \sum_{i=1}^{n} \mathrm{Var}(X_i)$$

$$= \sum_{i=1}^{n} \frac{35}{12} = 10 \cdot \frac{35}{12}$$

What about the standard deviation?

Var has grown by factor 10,

σ grows by factor $\sqrt{10}$!