

Section 3.3. [cont'd].

Problem. Let  $X$  and  $Y$  be two random variables on the same  $\Omega$ .  
Let

$$\mu_X = \mathbb{E}[X] = 7, \quad \sigma_X = 1, \quad \rho_{X,Y} = \frac{\text{Cov}[X,Y]}{\sigma_X \cdot \sigma_Y}$$

$$\mu_Y = \mathbb{E}[Y] = -3, \quad \sigma_Y = 2, \quad \rho_{X,Y} = \text{corr}[X,Y] = -0.5.$$

Find  $\text{Var}[2X+Y-5]$ .

$$\rightarrow \text{Var}[2X+Y-5] = \text{Var}[2X+Y] = \text{shift}$$

$$\begin{aligned} &= \text{Var}[2X] + 2\text{Cov}[2X,Y] + \text{Var}[Y] \\ &= 4 \cdot \text{Var}[X] + 2 \cdot 2 \cdot \text{Cov}[X,Y] + \text{Var}[Y] = \sigma_Y^2 \\ &= 4 \cdot 1 + 4 \cdot (-0.5) \cdot 1 \cdot 2 + 2^2 \\ &= 4 - 4 + 4 = 4 \end{aligned}$$

□

Square Root Law.

Consider  $\{X_1, X_2, \dots, X_n, \dots\}$  a sequence of independent, identically distributed random variables. iid

Let  $S_n = X_1 + X_2 + \dots + X_n$

$$\bar{X}_n = \frac{1}{n} \cdot S_n$$

for all  $n = 1, 2, \dots$

Focus on  $S_n$ :

$$\begin{aligned} \cdot \mathbb{E}[S_n] &= \mathbb{E}[X_1 + X_2 + \dots + X_n] && \text{linearity of } \mathbb{E} \\ &= \mathbb{E}[X_1] + \mathbb{E}[X_2] + \dots + \mathbb{E}[X_n] && \text{identically distributed} \\ &= n \cdot \mu_X \end{aligned}$$

$$\begin{aligned} \cdot \text{Var}[S_n] &= \text{Var}[X_1 + X_2 + \dots + X_n] && \text{independence} \\ &= \text{Var}[X_1] + \text{Var}[X_2] + \dots + \text{Var}[X_n] && \text{identically distributed} \\ &= n \cdot \sigma_X^2 && \Rightarrow \text{SD}[S_n] = \sigma_X \sqrt{n} \end{aligned}$$

- Focus on  $\bar{X}_n$ :
- $\mathbb{E}[\bar{X}_n] = \mathbb{E}\left[\frac{1}{n} \cdot S_n\right] = \frac{1}{n} \cdot \mathbb{E}[S_n] = \frac{1}{n} \cdot n \cdot \mu_x = \mu_x$
  - $\text{Var}[\bar{X}_n] = \text{Var}\left[\frac{1}{n} S_n\right] = \frac{1}{n^2} \text{Var}[S_n]$   
 $= \frac{1}{n^2} \cdot n \cdot \sigma_x^2 = \frac{\sigma_x^2}{n}$
  - $\text{SD}[\bar{X}_n] = \frac{\sigma_x}{\sqrt{n}}$
- 

## Law of Averages (aka The Law of Large Numbers (LLN))

$$\frac{x_1 + x_2 + \dots + x_n}{n} \xrightarrow{n \rightarrow \infty} \mu_x$$

## The Normal Approximation Theorem.

(aka The Central Limit Theorem (CLT))

$$\frac{S_n - n \cdot \mu_x}{\sigma_x \sqrt{n}} \xrightarrow[n \rightarrow \infty]{D} N(0,1)$$

Then,  $\mathbb{P}[c \leq S_n \leq d] = \mathbb{P}\left[\frac{c - n \cdot \mu_x}{\sigma_x \sqrt{n}} \leq \frac{S_n - n \cdot \mu_x}{\sigma_x \sqrt{n}} \leq \frac{d - n \cdot \mu_x}{\sigma_x \sqrt{n}}\right]$

$\approx \mathbb{P}\left[\frac{c - n \cdot \mu_x}{\sigma_x \sqrt{n}} \leq Z \leq \frac{d - n \cdot \mu_x}{\sigma_x \sqrt{n}}\right]$

$= \Phi\left(\frac{d - n \cdot \mu_x}{\sigma_x \sqrt{n}}\right) - \Phi\left(\frac{c - n \cdot \mu_x}{\sigma_x \sqrt{n}}\right)$

Remark. We must use the continuity correction for r.v.s whose support contains consecutive integers.

### Problem 3.3.24. from Pitman.

A box contains 4 tickets numbered 0, 1, 1, 2.

Let  $S_n$  be the sum of the numbers obtained from  $n$  draws w/ replacement @ random.

A single draw  $X$  has the pmf

$x$	0	1	2
$P_X(x)$	1/4	1/2	1/4

The sum  $S_2$  of two independent draws has the pmf

$s$	0	1	2	3	4
$P_{S_2}(s)$	1/16	1/4	3/8	1/4	1/16

$$S_2 = 1 \text{ iff } X_1 = 0 \text{ and } X_2 = 1 \quad \frac{1}{4} \cdot \frac{1}{2}$$

OR

$$X_1 = 1 \text{ and } X_2 = 0 \quad + \quad \frac{1}{2} \cdot \frac{1}{4}$$


---


$$\frac{1}{4}$$

$S_2 = 3$  same by symmetry

Using the CLT, approximate

$$P[S_{50} = 50]$$

$$\text{We need } \therefore \mu_X = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$$

$$\cdot \sigma_X = \sqrt{\text{Var}[X]}$$

$$\text{Var}[X] = E[X^2] - (E[X])^2$$

$$E[X^2] = 0^2 \cdot \frac{1}{4} + 1^2 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{4} = \frac{3}{2}$$

$$\sigma_X = \sqrt{\frac{3}{2} - 1} = \frac{\sqrt{2}}{2}$$

Symmetry of  $N(0,1)$

$$1 - \Phi(0.1)$$

We do need the continuity correction!

$$\Phi\left(\frac{50.5 - 50}{\frac{\sqrt{2}}{2} \cdot \sqrt{50}}\right) - \Phi\left(\frac{49.5 - 50}{\frac{\sqrt{2}}{2} \cdot \sqrt{50}}\right) = \Phi(0.10) - \Phi(-0.10)$$

$$= 2 \cdot \Phi(0.10) - 1$$

$$= 2 \cdot 0.5398 - 1 = 0.0796$$

□