(1) Prove equation 5.7:

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

Start w/ our definition of variance:

$$Var(X) = E([X - E(X)]^2) \qquad \text{eq. 5.6}$$

Squaring the inside expression:

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

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

Using linearity of expectation:

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

E(X) is a constant

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

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

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

$$= E(X^2) - [E(X)]^2$$ -> finished: see eq. 5.7

(2) Prove $Var(a + cX) = c^2 Var(X)$ in 2 steps

(2a) Prove $Var(X + c) = Var(X)$

Hint: Start with

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

$$Var(X+c) = E\left(\left[X + c - \boxed{E(X+c)}\right]^2\right)$$

Linearity of expectation

$$= E\left(\left[X + c - \left(E(X) + c\right)\right]^2\right)$$

$$= E\left(\left[X + \cancel{c} - E(X) - \cancel{c}\right]^2\right)$$

$$= \underbrace{E\left(\left[X - E(X)\right]^2\right)}_{\text{this is } Var(X)}$$

(2b) Prove that $\mathrm{Var}(cX) = c^2\mathrm{Var}(X)$

Hint: start w/ 5.7

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

$$\mathrm{Var}(cX) = E([cX]^2) - [E(cX)]^2$$

$$= c^2E(X^2) - [cE(X)]^2$$

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

$$= c^2\left[\boxed{E(X^2) - [E(X)]^2}\right]$$

$$= c^2\mathrm{Var}(X)$$

(3). Prove that $Var(X+Y) = Var(X) + Var(Y)$

(3a) assume X and Y are discrete random variables.

we need the joint pmf of X and Y:

$$f_{X,Y}(x,y) = P(X=x \cap Y=y)$$

i.e. the probability that $X=x$ and $Y=y$

what is $f_{X,Y}(x,y)$ in terms of $f_X(x)$ and $f_Y(y)$?

If f_X and f_Y are independent, then we can multiply these probabilities.

so:

$$f_{X,Y}(x,y) = P(X=x \cap Y=y)$$

$$= P(X=x) \times P(Y=y)$$

$$= \underbrace{f_X(x)\, f_Y(y)}_{\text{"the product of the marginal mass functions of X and Y"}}$$

marginal : ignoring any other random variables

(3b) Express $Var(X+Y)$ in terms of $Var(X)$, $Var(Y)$, and the quantity $Cov(X,Y) = E(XY) - E(X)E(Y)$

It would be useful to give a hint:

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

$$Var(X+Y) = E\left(\underbrace{(X+Y)^2}\right) - \underbrace{\left[E(X+Y)\right]^2}$$

$$E[(X+Y)(X+Y)]$$

$$E[X^2 + 2XY + Y^2]$$

$$= E(X^2) + E(2XY) + E(Y^2)$$

$$= E(X^2) + 2E(XY) + E(Y^2)$$

$$[E(X+Y)][E(X+Y)]$$

$$= [E(X) + E(Y)][E(X) + E(Y)]$$

$$= [E(X)]^2 + E(X)E(Y) + E(Y)E(X) + [E(Y)]^2$$

$$= [E(X)]^2 + 2E(X)E(Y) + [E(Y)]^2$$

$$\Downarrow$$

$$= E(X^2) + 2E(XY) + E(Y^2) - [E(X)]^2 - 2E(X)E(Y) - [E(Y)]^2$$

$$= \boxed{E(X^2) - [E(X)]^2} + \boxed{E(Y^2) - [E(Y)]^2} + \boxed{2E(XY) - 2E(X)E(Y)}$$

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

(3c) $E(XY) = \sum_x \sum_y xy f_{X,Y}(x,y)$

Prove that when X and Y are independent,

$$E(XY) = E(X)E(Y)$$

a reminder : $E(X) = \sum x_i f_X(x_i)$

If X and Y are independent, then

$$f_{X,Y}(x,y) = f_X(x) f_Y(y)$$ } from (3a)

So:

$$E(XY) = \sum_x \sum_y xy \, f_X(x) f_Y(y)$$

Rearranging..

$$= \boxed{\sum_x x f_X(x)} \boxed{\sum_y y f_Y(y)}$$

$$= \boxed{E(X)} \boxed{E(Y)}$$

What does this imply about $Var(X+Y)$ when X and Y are independent?

$$Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y) \quad \text{(from 3b)}$$

$$= Var(X) + Var(Y) + 2\left[E(X,Y) - E(X)E(Y)\right]$$

$$= Var(X) + Var(Y) + \boxed{2E(X,Y)} - 2E(X)E(Y)$$

$$= Var(X) + Var(Y) + \cancel{2\boxed{E(X)E(Y)}} - \cancel{2E(X)E(Y)}$$

from 3c : independence of X and Y

$$Var(X+Y) = Var(X) + Var(Y)$$

(3d) What is $Var(X-Y)$ when X and Y are independent?

Notice that $X - Y = X + cY$, where $c = -1$

$$Var(X+cY) = Var(X) + Var(cY)$$

$$Var(cY) = c^2 Var(Y)$$

$$= Var(X) + c^2 Var(Y)$$

$$= Var(X) + (-1)^2 Var(Y)$$

$$Var(X-Y) = Var(X) + Var(Y)$$

(4a) What is the variance of a Bernoulli random variable w/ parameter p?

$$P(X=k) = p^k (1-p)^{1-k}$$

$$E(X) = p$$

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

$$\boxed{E(X^2) = E(X) = p}$$

why? Because for the Bernoulli,

$$0^2 = 0$$
$$1^2 = 1$$

$$= p - p^2$$
$$= p(1-p)$$

$$E[P(X=k)] = 0 \cdot p^0 (1-p)^{(1-0)} + 1 \cdot p^1 (1-p)^{(1-1)}$$
$$= 0 + p$$
$$= p$$

(4b) Use part (a) and the result from problem 3 to get the variance of a binomial random variable.

$$Var(X+Y) = Var(X) + Var(Y)$$

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$E(X) = np$$

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

Binomial

Bernoulli

Here we express the Binomial as the sum of n Bernoulli trials

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

$$= \sum_{i=1}^{n} \boxed{\mathrm{Var}(X_i)}$$

$$= \sum_{i=1}^{n} \boxed{p(1-p)}$$

$$= np(1-p)$$

(5) What is the variance of $\frac{1}{n}(X_1 + X_2 + \dots X_n)$, the mean of the X_i?

$$\mathrm{Var}\left(\frac{1}{n}(X_1 + X_2 + \dots + X_n)\right)$$

$\mathrm{Var}(cX) = c^2 \mathrm{Var}(X)$, so:

$$\frac{1}{n}^2 \mathrm{Var}(X_1 + X_2 + \dots + X_n)$$

$$\frac{1}{n^2} \underbrace{Var\left(X_i \ldots\right)}$$

Variance of the sums is equal to the sum of the variances (property of variances)

$$= \frac{1}{n^2} \sum_{i=1}^{n} \theta^2$$

$$= \frac{1}{\cancel{n^2}} \cancel{n} \, \theta^2$$

$$= \frac{1}{n} \theta^2$$ Variance of the mean of X_i's

$$\sqrt{\frac{1}{n} \theta^2}$$ SD = $\sqrt{\text{variance}}$

$$= \frac{\theta}{\sqrt{n}}$$: SD of the mean of X_i's

= <u>standard error of the mean SEM</u>

(6) Proving one version of the law of large numbers - the "weak" version.

(6a) Proving Markov's inequality: for a non-negative random variable X:

$$P(X \geq c) \leq \frac{E(X)}{c}$$

where c is a positive constant.

$$Z = \begin{cases} X \geq c, & Z = 1 \\ X < c, & Z = 0 \end{cases}$$

We define Z, a Bernoulli random variable

↳ Bernoulli, w/ $p = P(X \geq c)$

Expectation of a Bernoulli random variable is p, so:

$$P(X \geq c) = E(Z)$$

If $X \geq c$, then $\frac{X}{c} \geq 1$ and $Z = 1$ (b/c Bernoulli),

so $Z \leq \frac{X}{c}$

$\left(\text{b/c } \frac{X}{c} \geq Z \text{ is = to } Z \leq \frac{X}{c} \right)$

$$\left(\begin{aligned} X &\geq c \\ \frac{X}{c} &\geq \frac{c}{c} \\ \frac{X}{c} &\geq 1 \end{aligned} \right)$$

Divide both sides by c

If $x < c$, and $Z = 0$,

so $\frac{X}{c} < 1$ $\left(1 > \frac{X}{c} \right)$

we also realize again that $Z \leq \frac{X}{c}$, because X is non-negative, and c is positive (and $Z = 0$)

If $Z \leq \frac{X}{c}$ in all cases, then:

$$E(Z) \leq E\left(\frac{X}{c}\right)$$

$$E(Z) \leq \frac{1}{c} E(X)$$

$$E(Z) \leq \frac{E(X)}{c}$$

$$P(X \geq c) \leq \frac{E(X)}{c}$$ (Markov's inequality)

(6b) Chebyshev's inequality says that for a random variable Y, with expectation μ,

$$P(|Y-\mu| \geq d) \leq \frac{Var(Y)}{d^2}$$

where d is a positive constant

Prove chebyshev's inequality by replacing X in Markov's inequality with $(Y-\mu)^2$

$$P(X \geq c) \leq \frac{E(X)}{c}$$ Markov's inequality

$$P(\underbrace{(Y-\mu)^2}_{\text{this is non-negative, so OK}} \geq c) \leq \frac{\overbrace{E((Y-\mu)^2)}^{\text{Definition of Var}(Y)}}{c}$$

$$P((Y-\mu)^2 \geq c) \leq \frac{Var(Y)}{c}$$

Now define another constant d as $d = \sqrt{c}$

This means that

$$P\left(\sqrt{(Y-\mu)^2} \geq \sqrt{c}\right) \leq \frac{Var(Y)}{c}$$

OR

$$P(|Y-\mu| \geq d) \leq \frac{Var(Y)}{d^2}$$ $d^2 = c$

(6c) $E(X_1) = E(X_2) = \dots = E(X_n) = \mu$

$\mu < \infty$

Define $\bar{X}_n$ as:

$$\bar{X}_n = \frac{1}{n}(X_1 + X_2 + X_3 + \dots + X_n)$$

For any positive constant δ, $\lim_{n \to \infty} P(|\bar{X}_n - \mu| > \delta|) = 0$

Use chebyshev's inequality: $P(|Y-\mu| \geq d) \leq \frac{Var(Y)}{d^2}$

and the variance of X_i is $= \frac{1}{n}\sigma^2$

to prove the 'weak' law.

Also, $Var(X_n) = \sigma^2$ and $\sigma^2 < \infty$

$$P(|\bar{X}_n - \mu| > \delta) \leq \frac{Var(\bar{X}_n)}{\delta^2}$$ chebyshev's inequality

$$\leq \frac{\frac{1}{n}\sigma^2}{\delta^2}$$

as $n \to \infty$, numerator goes to 0

$$P\left(|\bar{X}_n - \mu| > \delta\right) \quad \boxed{\leq} \; 0$$

can't be negative

$$P\left(|\bar{X}_n - \mu| > \delta\right) = 0$$