Ashwin

## 3.3 Expected values

Consider a university with 15'000 students. Let $X$ be the number of courses for which a random student is registered. (Assume) The pf of $X$ is:

| $x$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $p(x_n)$ | 0.01 | 0.03 | 0.13 | 0.25 | 0.39 | 0.17 | 0.02 |
| # of students | $0.01 \times 15'000 = 150$ | 450 | 1950 | 3750 | 5850 | 2550 | 300 |

The average number of courses per student is

$$\frac{1 \cdot 150 + 2 \cdot 450 + \ldots + 7 \cdot 300}{15'000}$$

$$= 1 \cdot 0.01 + 2 \cdot 0.03 + \ldots + 7 \cdot 0.02 = 4.57.$$

**Rem:** To compute the average of $X$ over the population, we did not have to use the population size. Knowledge of the pf was enough.

**Def.** Let $X$ be a discrete rv with pf $p$. The *expected value* (or *mean value*) of $X$ is

$$\mathbb{E}X := \sum_{x \in \mathbb{R}} x \cdot p(x).$$

f

Ex: i. For the above example,

$$\mathbb{E}X = \sum x p(x) = 1 \cdot p(1) + 2p(2) + \dots + 7p(7)$$
$$= 1 \cdot 0.01 + 2 \cdot 0.03 + \dots + 7 \cdot 0.02$$
$$= 4.57.$$

ii. Let $X$ be a Bernoulli rv with pf $p(1) = p$ and $p(0) = 1-p$, where $p \in (0,1)$. Then,

$$\mathbb{E}X = \sum x p(x) = 0 \cdot p(0) + 1 \cdot p(1) = p.$$

iii. Let $X$ be a rv with pf

$$p(x) := \begin{cases} k/x^2 & , x = 1, 2, \dots \\ 0 & , \text{else}. \end{cases}$$

Then, $k = \frac{\pi^2}{6}$ ensures that $\sum_{x \in \mathbb{N}} p(x) = 1$. (You don't need to understand why. It is not easy.) The expected value of $X$ is

$$\mathbb{E}X = \sum_x x p(x) = k \sum_{x \in \mathbb{N}} \frac{1}{x}. \quad \text{"harmonic series"}$$
$$= \infty.$$

So, when $p(x)$ does not decrease fast enough as $x$ increases, then $X$ does not have an finite expectation. We say $X$ is heavy-tailed.

## The expected value of a function

Suppose the cost of a diagnostic vehicle test depends on the number of cylinders $X$ in the vehicle via $h(X) = 20 + 3X + 0.5X^2$. If $X$ is a rv, then so is $Y = h(X)$. Suppose the pf of $X$ is:

| $x$. | 4 | 6 | 8 |
|---|---|---|---|
| $p(x)$. | 0.5 | 0.3 | 0.2 |

$\Longrightarrow$

| $y$. | 40 | 56 | 76 |
|---|---|---|---|
| $p(y.)$ | 0.5 | 0.3 | 0.2 |

$$\Longrightarrow \mathbb{E}Y = \sum_y y\,p(y) = 40 \cdot 0.5 + 56 \cdot 0.3 + 76 \cdot 0.2$$

$$= h(4) \cdot 0.5 + h(6) \cdot 0.3 + h(8) \cdot 0.2$$

$$= \sum_x h(x)\,p(x).$$

**Prop:** Let $X$ be a rv with pf $p$. Then,

$$\mathbb{E}\,h(X) = \sum_{x \in \mathbb{R}} h(x)\,p(x). \qquad [!?]$$

**Example:** A store purchased three computers at 500 \$ a piece. It will try to sell them at 1000 \$ each. If it cannot it will return them to the manufacturer who will refund 200 \$ to the store. What is the expected profit of the store if the pf of the number of computers sold $X$ is

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $p(x)$ | 0.1 | 0.2 | 0.3 | 0.4 |

The profit is $h(X) = 1000X + 200(3-X) - 1500$. So,

$$\mathbb{E}h(X) = \sum h(x)p(x) = h(0)p(0) + h(1)p(1) + h(2)p(2) + h(3)p(3)$$

$$= -900 \cdot 0.1 + (-100) \cdot 0.2 + 700 \cdot 0.3 + 1500 \cdot 0.4$$

$$= 700.$$

## The expected value of a linear function

**Prop:** $\mathbb{E}(aX+b) = a\mathbb{E}X + b$. "Linearity of expectation"

**Proof:** $\mathbb{E}(aX+b) = \sum (ax+b)p(x) = a\sum xp(x) + b\sum p(x)$

$$= a\mathbb{E}X + b$$

**Cor:** i $\mathbb{E}(aX) = a\mathbb{E}X$ ii $\mathbb{E}(X+b) = \mathbb{E}X + b$.

**Example:** In the previous example $h(X) = 800X - 900$ and $\mathbb{E}X = 0.2 + 0.6 + 1.2 = 2.0$

$$\Rightarrow \mathbb{E}h(X) = 800\,\mathbb{E}X - 900 = 700.$$

## The variance of X

The expected value of $X$ describes where the distr. of $X$ is centered. The variance will describe its spread or variability:

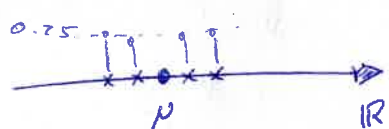
**Def:** Let $X$ be a discrete rv. with pf $p$ and mean value $\mu$. The variance of $X$ is

$$\sigma_X^2 = \sigma^2 = V(X) = \sum_{x \in \mathbb{R}} (x-\mu)^2 p(x) = \mathbb{E}((X-\mu)^2).$$

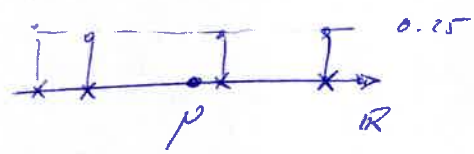
The standard deviation (sd) of $X$ is

$$\sigma_X = \sqrt{\sigma_X^2}.$$

**Rem:** $(X-\mu)^2$ is the squared deviation of $X$ from its mean st $\sigma_X^2$ is the expected squared deviation.




small variance / sd




large variance / sd.

**Ex** Consider a discrete rv $X$ with pf

| $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $p(x)$ | 0.3 | 0.25 | 0.15 | 0.05 | 0.1 | 0.15 |

The expected value of $X$ is

$$\mathbb{E}X = \sum x p(x) = 0.3 + 0.5 + \cdots + 0.9 = 2.85$$

The variance of $X$ is

$$V(X) = \sum (x-2.85)^2 p(x) = (1-2.85)^2 \cdot 0.3 + (2-2.85)^2 \cdot 0.25 + \dots + (6-2.85)^2 \cdot 0.15$$

$$= \dots = 3.2275$$

$$\Rightarrow \quad \sigma_X = \sqrt{3.2275} \simeq 1.800.$$

Prop: $V(X) = \sum_{x \in \mathbb{R}} x^2 p(x) - \mu^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2$ !

Proof: $\sigma^2 = \mathbb{E}((X-\mu)^2) = \mathbb{E}(X^2 - 2\mu X + \mu^2)$

$$= \mathbb{E}(X^2) - 2\mu \mathbb{E}X + \mu^2 = \mathbb{E}(X^2) - 2\mu^2 + \mu^2 = \mathbb{E}(X^2) - \mu^2 \quad \square$$

Example: In the example above:

$$\mathbb{E}(X^2) = \sum x^2 p(x) = 1 \cdot 0.3 + 4 \cdot 0.25 + \dots + 36 \cdot 0.15$$
$$= 11.35.$$

$$\longrightarrow \quad \sigma^2 = \mathbb{E}(X^2) - \mu^2 = 11.35 - (2.85)^2 = 3.2275.$$

Variance of a linear function: $\longrightarrow$ Lecture 9

We have that

$$V(h(X)) = \sum_x (h(x) - \mathbb{E}(h(X)))^2 p(x).$$

If $h(X) = aX + b$, then

$$h(X) - \mathbb{E}(h(X)) = aX + b - a\mu - b = a(X - \mu).$$

**Prop:** $V(aX+b) = a^2 V(X)$, $\sigma_{aX+b} = |a|\sigma_X$ [Pr]

**Cor:** ~~i.~~ i. $\sigma_{aX} = |a|\sigma_X$, ii. $\sigma_{X+b} = \sigma_X$.

**Rem:** Please, don't forget the absolute value!

## 3.4 The binomial probability distr.

Many experiments conform exactly or approximately to the following requirements:

i. The experiment consists of a sequence of $n$ other experiments, called trials.

ii. Each trial has two possible outcomes. Let's say $S$ and $F$. → labels are arbitrary

iii. The trials are independent: the outcome of any trial does not influence the outcome of any other trial.

iv. $\mathbb{P}(S)$ is constant from trial to trial.

We call such an example a binomial experiment.

**Ex.** i. Tossing a coin $n \in \mathbb{N}$ times is a binomial experiment.

ii. Drawing $n$ balls from an urn containing $k$ blue ones and $\ell$ red ones with replacement, is a binomial experiment.

Replace the ball after every draw.

$$\mathbb{P}(\text{a blue ball is drawn}) = \frac{k}{k+\ell}$$

$$\mathbb{P}(\text{a red ball is drawn}) = \frac{\ell}{k+\ell}$$

iii. Think of drawing 6 balls without replacement from an urn containing 35 blue ones and 15 red ones. Then,

i. $\mathbb{P}(B \text{ on second draw} \mid B \text{ on first draw})$

$$= \frac{34}{49} \cong 0.6939$$

ii. $\mathbb{P}(B \text{ on second draw})$

$$= \mathbb{P}(B \text{ on second} \mid B \text{ on first})\,\mathbb{P}(B \text{ on first}) + \mathbb{P}(B \text{ on second} \mid R \text{ on first})\,\mathbb{P}(R \text{ on first})$$

$$= \frac{34}{49}\cdot\frac{35}{50} + \frac{35}{49}\cdot\frac{15}{50} = \frac{35}{50}\left(\frac{34}{49}+\frac{15}{49}\right)$$

$$= \frac{35}{50} \cong 0.7.$$

So, the trials are <u>not</u> independent. This can also be seen from

$$\mathbb{P}(B \text{ on sixth} \mid BBBBB) = \frac{30}{45} = \frac{2}{3} \doteq 0.\dot{6}$$

$$\mathbb{P}(B \text{ on sixth} \mid RRRRR) = \frac{35}{45} = \frac{7}{9} = 0.\dot{7}$$

iv. Now draw 6 balls without replacement from an urn containing $3.5 \cdot 10^6$ ~~red~~ blue balls and $1.5 \cdot 10^6$ red balls. Then,

$$\mathbb{P}(B \text{ on sixth} \mid BBBBB) = \frac{3'499'995}{4'999'995} \simeq 0.7000,$$

$$\mathbb{P}(B \text{ on sixth} \mid FFFFF) = \frac{3'500'000}{4'999'995} \simeq 0.7000.$$

So, here trials are effectively independent.

Upshot Sampling / Drawing with replacement always yields independent trials. Sampling / Drawing without replacement does so approximately if the population is very large. !

Def. i The binomial rv $X$ associated with a binomial experiment is

$$X := \# \text{ of } S \text{ among } n \text{ trials} \qquad [!?]$$

ii. The pf of a binomial rv depends on $n$ and on the probability $p$ of $S$: we denote it by $b(x; n, p)$.

Ex. Consider 20 male and 10 female Monarch butterflies. Pick four butterflies with replacement. ~~What is the probability that~~ Let $X$ be the number of male butterflies picked. What is the pf of $X$?

$X$ can take the values $0, 1, 2, 3, 4$:

$$\mathbb{P}(X=0) = \mathbb{P}(FFFF) = \mathbb{P}(F)^4 = \left(\frac{10}{10+20}\right)^4 = \frac{1}{3^4}$$

$$\mathbb{P}(X=1) = \mathbb{P}(MFFF \text{ or } FMFF \text{ or } FFMF \text{ or } FFFM)$$
$$= 4\, \mathbb{P}(M)\, \mathbb{P}(F)^3 = 4\, \frac{2}{3} \frac{1}{3^3}$$

$$\mathbb{P}(X=2) = \binom{4}{2} \mathbb{P}(M)^2\, \mathbb{P}(F)^2 = \binom{4}{2} \frac{2^2}{3^2} \frac{1}{3^2}$$

$$\mathbb{P}(X=3) = \binom{4}{3} \frac{2^3}{3^3} \frac{1}{3}$$

$$\mathbb{P}(X=4) = \binom{4}{4} \frac{2^4}{3^4}.$$

The above is an experiment with replacement; i.e., a binomial experiment. So,

$$\mathbb{P}(X=n) = b(n;4,\tfrac{2}{3}) = \begin{cases} \binom{4}{n}\left(\frac{2}{3}\right)^n\left(\frac{1}{3}\right)^{4-n} & n=0,1,2,3,4, \\ 0 & \text{else}. \end{cases}$$

$\mathbb{P}(\mu)$

**Theorem:** $b(x;n,p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{if } x=0,1,\dots,n \\ 0 & \text{else} \end{cases}$

**Proof:** $b(x;n,p) = \mathbb{P}(X=x) = \left(\text{number of sequences of length } n \text{ consisting of } x \ S\text{'s}\right)$

$\times \left(\text{probability of any such sequence}\right)$

$$= \binom{n}{x} p^x (1-p)^{n-x}$$

probability of $x$ $S$'s  probability of $n-x$ $F$'s

**Def:** If $X$ has pf $b(x;n,p)$, we write $X \sim \text{Binom}(n,p)$.

The cdf of $X$ is

$$B(x;n,p) := \mathbb{P}(X \le x) = \sum_{y=0}^{x} b(y;n,p).$$

**Example:** Suppose 20% of all copies of a particular textbook fail a certain binding strength test. Let $X$ be the number of copies that fail the test among 15 randomly selected copies.

a. What is the prob. that at most 8 fail the test?

b. " " exactly " "

c. " " at least " "

d. What is the probability that between 4 and 7 (inclusive) fail the test?

a. ~~P(X≤~~ $X \sim \text{Binom}(15, 0.2)$

$$\mathbb{P}(X \leq 8) = B(8; 15, 0.2) = \sum_{x=0}^{8} b(x; 15, 0.2)$$

$$\cong 0.999$$ → calculator

b. $$P(X = 8) = b(8; 15, 0.2) = \binom{15}{8} 0.2^8 \, 0.8^7 \cong 0.0035$$

c. $$P(X \geq 8) = 1 - P(X < 8) = 1 - \mathbb{P}(X \leq 7)$$
$$= 1 - B(7; 15, 0.2) \cong 0.004.$$

d. $$P(X \in \{4,5,6,7\}) = P(X \leq 7) - P(X \leq 3)$$
$$= B(7; 15, 0.2) - B(3; 15, 0.2)$$
$$\cong 0.348.$$

Rem: Learn how to do this with your calculator or with Table A.1 in Devore's book.

→ Lecture 10