L31: Weds., Apr. 12.

Housekeeping: · Written HW due Friday – see Canvas

· Book project – final essay 1st draft due Apr. 19 at 11:59 p.m. on Canvas – also bring a hard copy to class for peer review on Apr. 21

· Extra credit on infographics – no due date, but e-mail me the day before you plan to present.

Last time: Mean & s.d. of PDF / random variable

Questions?

Today: Binomial distribution.

EXAMPLE. Consider an experiment ... Pea plants can have pods that are green (dominant) or yellow (recessive). Suppose that two heterozygous pea plants are crossed (mated), and their offspring are numerous – enough to fill several greenhouses.

The probability that a single offspring plant has yellow or green pods is computed using classical / Mendelian genetics:

	G	y
G	GG	Gy
y	Gy	yy

the probability of a single offspring having a green pod is: $P(\text{green}) = \frac{3}{4} = 0.75$.

(You won't be tested / quizzed on Punnett squares...)

But...

If you choose 5 offspring plants at random from the greenhouse, and define X to be the number of plants of those 5 that have green pods...

what values can X take on?

0, 1, 2, 3, 4, or/and 5.

How to compute $P(X=3)$? (That is, P(3 of 5 are green) = ?)

... First, an easier question:

How to compute $P(GGGYY)$?

(That is, P(the 1st 3 are green) = ?)

Key observation: If you have many plants, then ~~selecti~~ the color of your first selection has no effect on the color of your second selection, and so forth...

Rule of thumb: If the "sample size" (e.g., 5 plants) is ≤ 5% of the populati[on] si[ze] then we treat selec'ns as ind[ependent]

Therefore: ~~P~~(1st is green) and ~~P~~(2nd is green) are INDEPENDENT EVENTS.

Also, (2nd is green) and (3rd is green) are INDEPENDENT.

And (3rd is green) and (4th is yellow) are independent.

And (4th is yellow) and (5th is yellow) are independent.

So.... USE THE MULTIPLICATION RULE:

$$P(GGGYY) = P\big((1^{st}\text{ is grn}) \text{ AND } (2^{nd}\text{ is grn}) \text{ AND } (3^{rd}\text{ is grn}) \text{ AND } (4^{th}\text{ is ylw}) \text{ AND } (5^{th}\text{ is ylw})\big)$$

$$= P(1^{st}\text{ is grn}) \cdot P(2^{nd}\text{ is grn}) \cdot P(3^{rd}\text{ is grn}) \cdot P(4^{th}\text{ is ylw}) \cdot P(5^{th}\text{ is ylw})$$

$$= \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{1}{4} \cdot \frac{1}{4} = \frac{3^3}{4^5} \approx 0.026$$

So, $P(GGGYY) = \frac{3^3}{4^5}$.

Now... what's $P(GGYYG)$?

(think a moment)

$$P(GGYYG) = \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{1}{4} \cdot \frac{1}{4} \cdot \frac{3}{4} = \frac{3^3}{4^5}.$$

Multiplication is commutative!

In fact, ~~that~~ any way you make the arrangement of 3 greens & 2 yellows, the probability of THAT SPECIFIC ARRANGEMENT is $\frac{3^3}{4^5}$.

<u>Another key observation</u>: (GGGYY) and (GGYYG) cannot occur simultaneously — they are MUTUALLY EXCLUSIVE EVENTS.

So,

$$\begin{aligned} P(GGGYY \text{ or } GGYYG) &= P(GGGYY) + P(GGYYG) \\ &= \frac{3^3}{4^5} + \frac{3^3}{4^5} = 2\left(\frac{3^3}{4^5}\right). \end{aligned}$$

≈ 0.053.

So, the last piece of the puzzle:

$$P(X=3) = \begin{pmatrix}\text{\# of possible} \\ \text{arrangements of} \\ \text{3G and 2Y}\end{pmatrix} \cdot \left(\frac{3^3}{4^5}\right).$$

GGGYY
GGYYG
GG~~GYGY~~
YGGGY
⋮
$\}$ # ?

So, what is the # of possible arrangements of 3G & 2Y?

That is... there are 5 pea pods — in how many ways can you choose 3 of them to be green?

$$_5C_3 = \frac{5!}{(5-3)!\,3!} = \frac{5 \cdot \overset{2}{\cancel{4}} \cdot \cancel{3 \cdot 2 \cdot 1}}{\cancel{(2\cdot 1)(3\cdot 2\cdot 1)}} = 5 \cdot 2 = 10$$

So, finally,

$$\boxed{P(X=3) = {_5C_3} \cdot \left(\frac{3}{4}\right)^3 \left(\frac{1}{4}\right)^2}$$

$$P(X=3) = 10 \cdot \frac{3^3}{4^5} \approx 0.264.$$

We can do similar calculations for the other values of X:

$$P(X=1) = {_5C_1} \cdot \left(\frac{3}{4}\right)^1 \cdot \left(\frac{1}{4}\right)^4 = \frac{5!}{(5-1)!\,1!} \cdot \frac{3}{4^5} = \frac{5 \cdot 3}{4^5} \approx 0.015$$

$$P(X=0) = {_5C_0} \cdot \left(\frac{3}{4}\right)^0 \left(\frac{1}{4}\right)^5 = \frac{5!}{(5-0)!\,0!} \cdot \frac{1}{4^5} = \frac{1}{4^5} \approx 0.000976 \approx 0.001$$

$$P(X=2) = {_5C_2} \cdot \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right)^3 = \frac{5!}{(5-2)!\,2!} \cdot \frac{3^2}{4^5} = \frac{10 \cdot 3^2}{4^5} \approx 0.0878$$

$$P(X=4) = {_5C_4} \cdot \left(\frac{3}{4}\right)^4 \left(\frac{1}{4}\right)^1 = \frac{5!}{(5-4)!\,4!} \cdot \frac{3^4}{4^5} = \frac{5 \cdot 3^4}{4^5} \approx 0.396$$

$$P(X=5) = {}_5C_5 \cdot \left(\frac{3}{4}\right)^5 \cdot \left(\frac{1}{4}\right)^0 = \frac{5!}{(5-5)!\,5!} \cdot \frac{3^5}{4^5} \approx 0.237$$

So... the PDF:

x	$P(X=x)$
0	0.001
1	0.015
2	0.088
3	0.264
4	0.396
5	0.237

Question: Suppose 75% (or 3 out of 4) customers at a store are happy with the service at that store.

Suppose the manager randomly selects 5 customers on a given day. Let X be the # of happy ones, out of those 5, he encounters.

Construct a PDF for X.

X is, in this case, a binomial random variable — just like the pea plant experiment; so, the PDF is the same.

For the pea plants: sample size $= 5$

$$P(X=0) = {}_{5}C_{\boxed{0}} \cdot \left(\frac{3}{4}\right)^{\boxed{0}} \left(\frac{1}{4}\right)^{\boxed{5}}$$

prob. of green

$$= \frac{5!}{(5-0)!\,0!} \left(\frac{3}{4}\right)^{0} \left(\frac{1}{4}\right)^{5}$$

Recall: $0! = 1$

$$= \frac{5!}{\cancel{5!}\cdot\cancel{1}} \cancel{\left(\frac{3}{4}\right)^{0}} \left(\frac{1}{4}\right)^{5}$$

$$= \left(\frac{1}{4}\right)^{5} \approx 0.000976 \approx 0.001$$

$$P(X=1) = {}_{5}C_{1} \left(\frac{3}{4}\right)^{1} \left(\frac{1}{4}\right)^{4}$$

$$\left(\frac{a}{b}\right)^{c} = \frac{a^{c}}{b^{c}}$$

$$= \frac{5!}{(5-1)!\,1!} \left(\frac{3}{4}\right)^{1} \left(\frac{1}{4}\right)^{4}$$

$$= \frac{5!}{4!} \left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^{4} = \frac{\boxed{5}\cdot\cancel{4\cdot3\cdot2\cdot1}}{\cancel{4\cdot3\cdot2\cdot1}} \cdot \frac{\boxed{3}}{4} \cdot \frac{1^{4}}{4^{4}}$$

$$= \frac{5\cdot3}{4^{5}}$$

$$\approx 0.015$$