6.2 The Binomial Distribution

Binomial Random Variable

If an experiment has only two, or can be expressed as only two outcomes, the random variable is called binomial.

Ex Tossing a coin & getting "H": b/c {H, T}

Ex Tossing a die and getting a 3:
b/c the outcome is "3", or, "not 3"

success ↑ "3"   failure ↑ "not 3"

X =	"3"	"not 3"
$P(X)$ =	1/6	5/6

Binomial Distributions require the following

1. Fixed number of trials per experiment
2. Two possible outcomes
3. The prob. of "success" is constant for each trial
4. The trials are independent
5. A random variable X representing the number of successes that occur.

* We will let $\begin{cases} n = \text{\# trials} \\ p = \text{probability a trial is successful.} \\ 1-p = \text{Prob. trial is a failure.} \end{cases}$

Note: We may call "catching a disease" a success but only from the statisticians viewpoint.

Ex Fair coin tossed 10 times, let X = # Heads.

Q: Binomial? ✓ yes H, or not H.

$\begin{cases} P(H) = 0.5, \quad P(\sim H) = 1 - 0.5 = 0.5 \\ X = \text{\# of heads} \\ n = 10 \end{cases}$

↑ not

Ex five basketball players take a free throw. let X = number of F.T. successfull.

Q: Binomial? 1. Trials Fixed - yes. (1 toss each)

2. Two outcomes? - yes (make it or miss it)

3. P(success) = constant?

No: each player will have different skill level so $P(\text{success}) \neq$ const

Ex 5 red & 5 green cards are in a box. Three are drawn at random w/o replacement. Let X = number of red cards drawn.

Q: Binomial? 1. $n = 3$

2. red or not red - yes

3. P(green card) = const? No... Cards are replaced.

5% Rule

* In large pools if 5% of the populations or less, is sampled and we do "replace" the subjects name back into the pool, then we can treat the sampling as independent.

This means $p(\text{success}) = \text{const}.$

EX 100ppl, pick 10 for a questionaire. You tell the subjects picked not to respond if picked a second time. "Not replacing"

- Here trials are not independent

If we sample only 5 people then we can assume the trials will be independent.

EX 1000 people, pick 50 or less, then you may consider polling without replacement and so the results will be independent

Lets toss a coin 3 times and as $P(2H)$ if we Consider a biased coin $p(H) = 0.6$

EX $P(2H \text{ in } 3 \text{ tosses})$ possible outcomes

$= P(HHT \text{ or } HTH \text{ or } THH)$

$= P(HHT) + P(HTH) + P(THH)$

$= (0.6)(0.6)(1-0.6) + (0.6)(1-0.6)(0.6) + (1-0.6)(0.6)(0.6)$

$= (0.6)^2(0.4) + (0.6)^2(0.4) + (0.6)^2(0.4)$

$= 3(0.6)^2(0.4)$

$= \boxed{0.432}$ 43.2% chance 3 tossed yeilds exactly 2 Heads.

⊛ Formula for a Binomial Probability.

$$\boxed{P(X) = \underbrace{{}_nC_X}_{} \; \underbrace{p^x}_{\text{probability of } X \text{ successes}} \; \underbrace{(1-p)^{n-x}}_{\text{prob. of } n-x \text{ failures}}} \qquad q \equiv 1-p$$

How many ways can I chose X successes out of n trials

Note: stat trek. com has a binomial prob. calculator.

stattrek.com/online-calculator/binomial

EX Pew Research found that 30% of internet users share photos on Pinterest. Suppose a random draw of $n=15$ internet users is performed.

a) Find the probability that exactly 4 pple use Pinterest.

$$\begin{cases} n = 15 & \text{\# of trials} \\ p = 0.3 \\ q = 1 - 0.3 = 0.7 \end{cases}$$

an internet user is selected

Find $P(X = 4 \text{ successes}) = \boxed{{}_{15}C_{4}\,(0.3)^4\,(0.7)^{11}}$

15 trials, 4 successes, 11 failures.

$= 1365\,(0.0081)(0.01977)$

$= 0.219$ or $\boxed{21.9\%}$ I draw 4 Pintrest users.

b) Find the probability of drawing fewer than 3 Pintrest users.

$$P(2 \text{ or } 1 \text{ or } 0) = P(2) + P(1) + P(0)$$

$$= {}_{15}C_{2}\,(0.3)^2(0.7)^{13} + {}_{15}C_{1}\,(0.3)^1(0.7)^{14} + {}_{15}C_{0}\,(0.3)^0(0.7)^{15}$$

$= 105$

$$= 0.09156 + 0.0305 + 0.00474 = \boxed{0.127}$$

* If "n" gets past a certain value we will desire to use a calculator or website or a table.

Excel

- The tables are used by • 1st picking the total number of trials in the binomial experiment (n)
  - • 2nd pick the number (x) of successes you desire (row).
  - • 3rd pick the prob. of success you desire (column)
- The intersection between row & col is the probability P (x successes in n trials)

See next page for such.

- For P (less than x) sum all the values above the row.
- For P (more than x) sum all values below the row

{ if "less than or equal" then include the value in the row itself.

Likewise for "more than or equal to"

Example 6.15

Use a table to compute binomial probabilities

The Pew Research Center recently reported that approximately 30% of Internet users in the United States use the image sharing website Pinterest. Suppose a simple random sample of 15 Internet users is taken. Use the binomial probability distribution to find the following probabilities.

a. Find the probability that exactly five of the sampled people use Pinterest.
b. Find the probability that fewer than four of the people use Pinterest.
c. Find the probability that the number of sampled people who use Pinterest is between 6 and 8, inclusive.

Solution

a. We have $n = 15$, so we go to the section of Table A.1 that corresponds to $n = 15$. This is shown in Figure 6.5. We look at the column corresponding to $p = 0.30$. Now for each value in the column labeled "x," the number in the table is the probability $P(x)$. We therefore look in the row corresponding to $x = 5$. The probability that exactly five people own a tablet computer is $P(5) = 0.206$.

							p							
n	x	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95
15	0	0.463	0.206	0.035	0.013	0.005	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+
	1	0.366	0.343	0.132	0.067	0.031	0.005	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+
	2	0.135	0.267	0.231	0.156	0.092	0.022	0.003	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+
	3	0.031	0.129	0.250	0.225	0.170	0.063	0.014	0.002	0.000+	0.000+	0.000+	0.000+	0.000+
	4	0.005	0.043	0.188	0.225	0.219	0.127	0.042	0.007	0.001	0.000+	0.000+	0.000+	0.000+
	5	0.001	0.010	0.103	0.165	0.206	0.186	0.092	0.024	0.003	0.001	0.000+	0.000+	0.000+
	6	0.000+	0.002	0.043	0.092	0.147	0.207	0.153	0.061	0.012	0.003	0.001	0.000+	0.000+
	7	0.000+	0.000+	0.014	0.039	0.081	0.177	0.196	0.118	0.035	0.013	0.003	0.000+	0.000+
	8	0.000+	0.000+	0.003	0.013	0.035	0.118	0.196	0.177	0.081	0.039	0.014	0.000+	0.000+
	9	0.000+	0.000+	0.001	0.003	0.012	0.061	0.153	0.207	0.147	0.092	0.043	0.002	0.000+
	10	0.000+	0.000+	0.000+	0.001	0.003	0.024	0.092	0.186	0.206	0.165	0.103	0.010	0.001
	11	0.000+	0.000+	0.000+	0.000+	0.001	0.007	0.042	0.127	0.219	0.225	0.188	0.043	0.005
	12	0.000+	0.000+	0.000+	0.000+	0.000+	0.002	0.014	0.063	0.170	0.225	0.250	0.129	0.031
	13	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.003	0.022	0.092	0.156	0.231	0.267	0.135
	14	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.005	0.031	0.067	0.132	0.343	0.366
	15	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.005	0.013	0.035	0.206	0.463

Figure 6.5

EXPLAIN IT AGAIN

Answers using technology may differ: If you find *P*(Fewer than 4) using technology, your answer will be 0.297 rather than 0.298. This difference isn't large enough to matter.

b. $P(\text{Fewer than } 4) = P(0) + P(1) + P(2) + P(3)$. We find these probabilities in Table A.1 and add them. See Figure 6.6.

$$P(\text{Fewer than } 4) = 0.005 + 0.031 + 0.092 + 0.170$$
$$= 0.298$$

							p							
n	x	0.05	0.10	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.75	0.80	0.90	0.95
15	0	0.463	0.206	0.035	0.013	0.005	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+
	1	0.366	0.343	0.132	0.067	0.031	0.005	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+
	2	0.135	0.267	0.231	0.156	0.092	0.022	0.003	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+
	3	0.031	0.129	0.250	0.225	0.170	0.063	0.014	0.002	0.000+	0.000+	0.000+	0.000+	0.000+
	4	0.005	0.043	0.188	0.225	0.219	0.127	0.042	0.007	0.001	0.000+	0.000+	0.000+	0.000+
	5	0.001	0.010	0.103	0.165	0.206	0.186	0.092	0.024	0.003	0.001	0.000+	0.000+	0.000+
	6	0.000+	0.002	0.043	0.092	0.147	0.207	0.153	0.061	0.012	0.003	0.001	0.000+	0.000+
	7	0.000+	0.000+	0.014	0.039	0.081	0.177	0.196	0.118	0.035	0.013	0.003	0.000+	0.000+
	8	0.000+	0.000+	0.003	0.013	0.035	0.118	0.196	0.177	0.081	0.039	0.014	0.000+	0.000+
	9	0.000+	0.000+	0.001	0.003	0.012	0.061	0.153	0.207	0.147	0.092	0.043	0.002	0.000+
	10	0.000+	0.000+	0.000+	0.001	0.003	0.024	0.092	0.186	0.206	0.165	0.103	0.010	0.001
	11	0.000+	0.000+	0.000+	0.000+	0.001	0.007	0.042	0.127	0.219	0.225	0.188	0.043	0.005
	12	0.000+	0.000+	0.000+	0.000+	0.000+	0.002	0.014	0.063	0.170	0.225	0.250	0.129	0.031
	13	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.003	0.022	0.092	0.156	0.231	0.267	0.135
	14	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.005	0.031	0.067	0.132	0.343	0.366
	15	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.000+	0.005	0.013	0.035	0.206	0.463

Figure 6.6