Learning Objectives

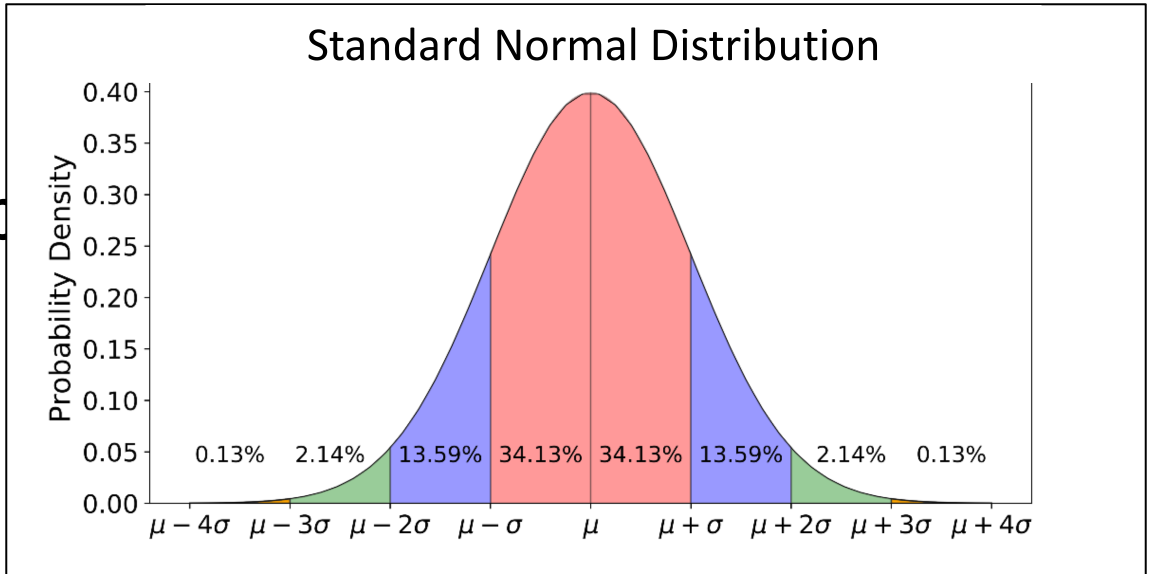
- **Review** inferential statistics, z-distribution, and models of randomness
- **Describe** 5 conditions required for using the *binomial distribution*
- **Visualize** probability in the binomial distribution
- **Solve** for expected probability using the binomial table

Review

- What is our goal when using *inferential statistics*?
- What does the *z*-distribution tell us?

Allows us to calculate a raw score for any given z-score, its percentile, and the probability of getting that score (from area under curve)

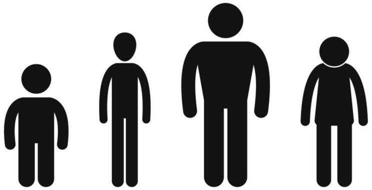
- What *z*-
distrib



Binomial Distribution

- Binomial distribution shows how likely observations are, when…
  - – 2 outcomes are possible
    - Example: Flipping a coin

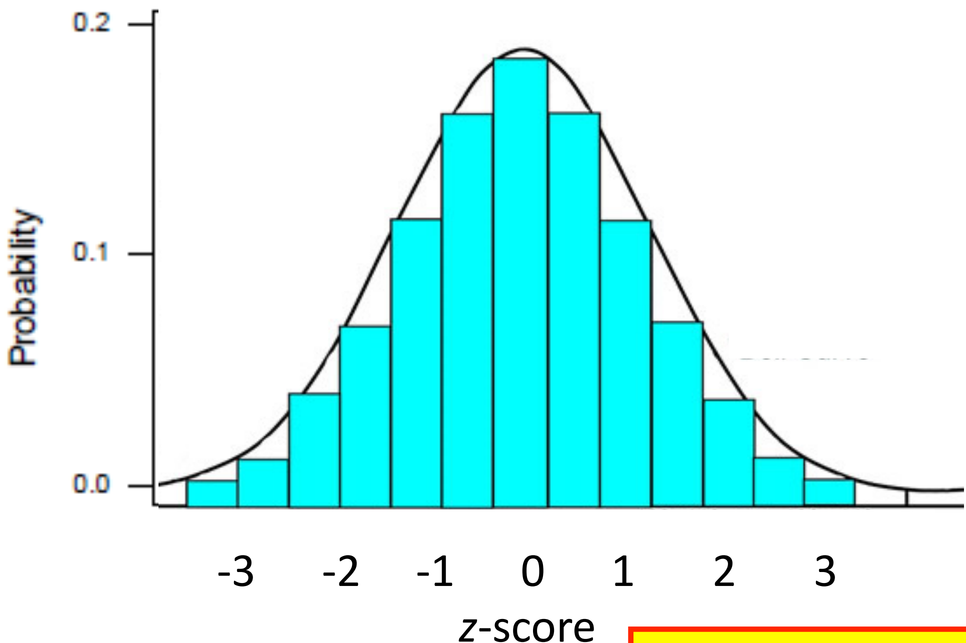
- *z*-distribution (or *normal distribution*) shows how likely observations are, when…
  - – Many outcomes are possible
    - Example: Heights relative to a population



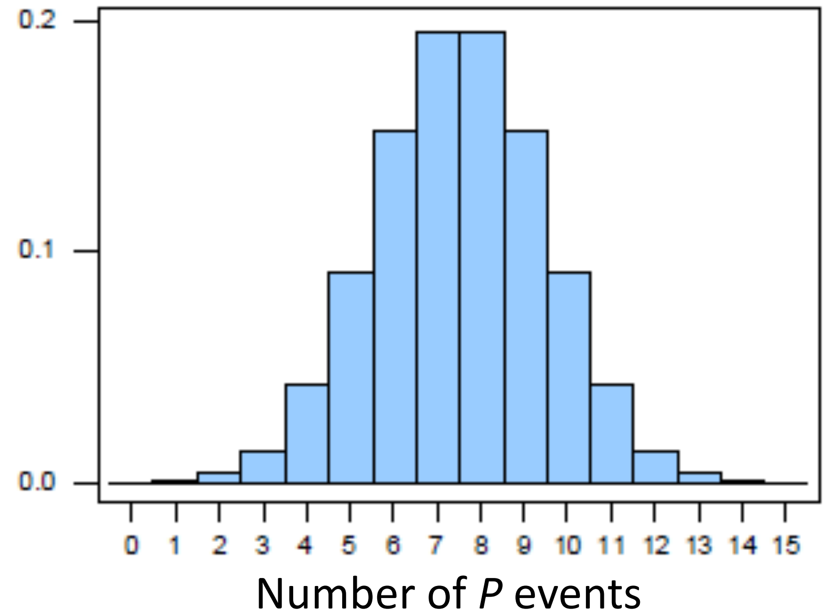
Modeling Probability

tallying probability for any number of P events

z-distribution



Binomial distribution



simplified version of z (practically indistinguishable graph)

Height of z=3 (or 6'8")?
That's RARE!!

14 out of 15 land on heads?
That's RARE!!

Test Yourself!

- A fair coin is flipped 3 times; what is the probability of getting <u>exactly 2 heads</u>?
    a) 0.667
    b) 0.375
    c) 0.250
    d) 0.125

- How many ways can we get 2 heads?
  - There are 3 different ways:
    1. HHT, p = .125
    2. HTH, p = .125
    3. <u>THH, p = .125</u>

    Sum(p's) = **.375**

Binomial Distribution

When $N = 3$ & $P = .50$

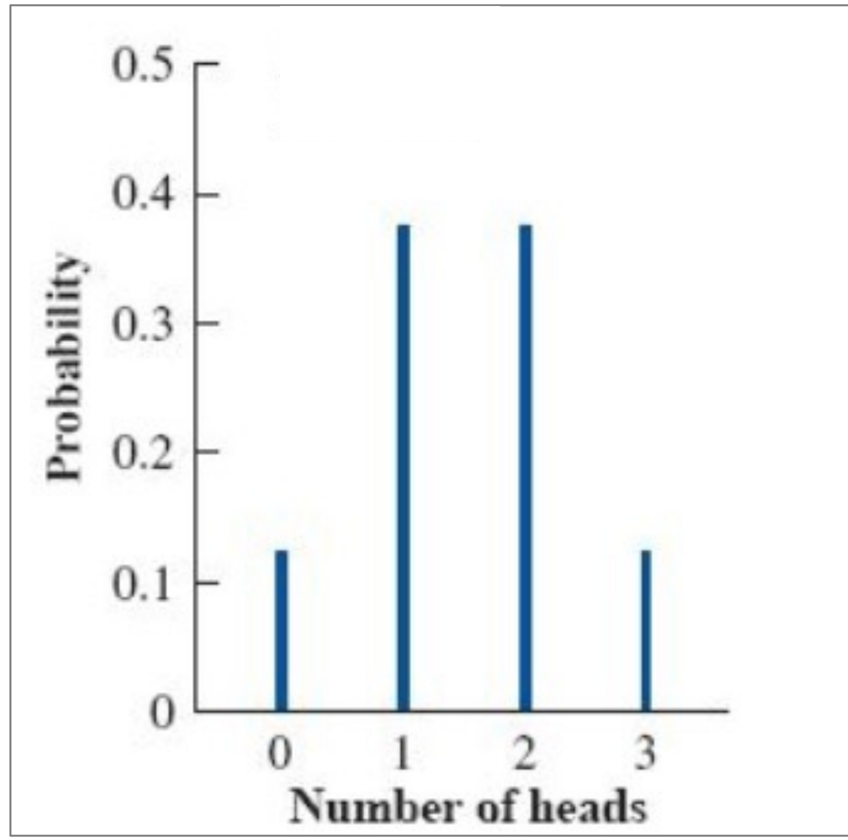
- p(2 heads)?

- p(0 heads)?

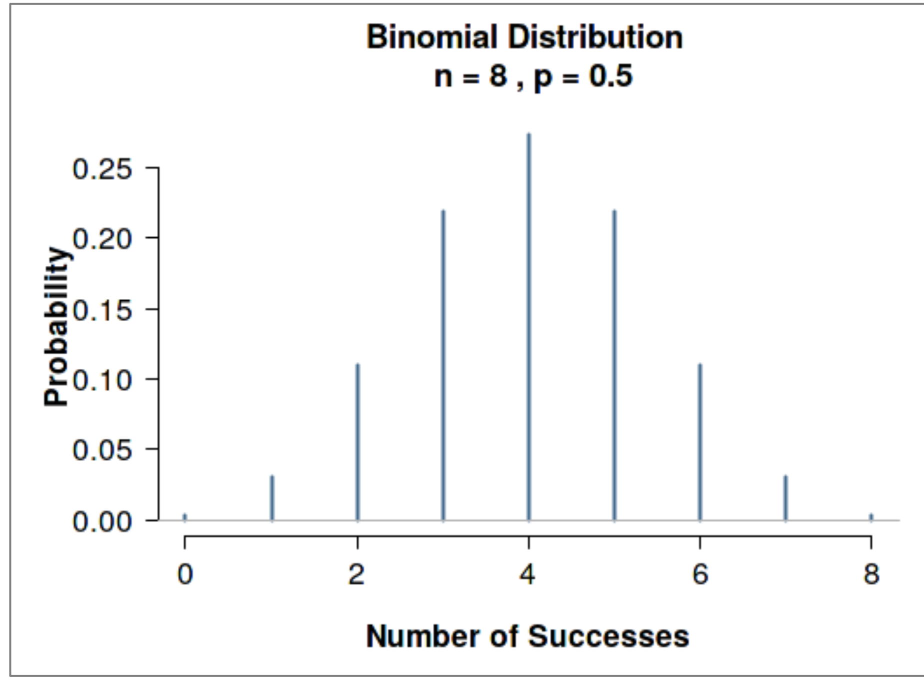
- p(*at least* 1 head)?

Need 2 pieces of info:

- N
- Expectation for P event

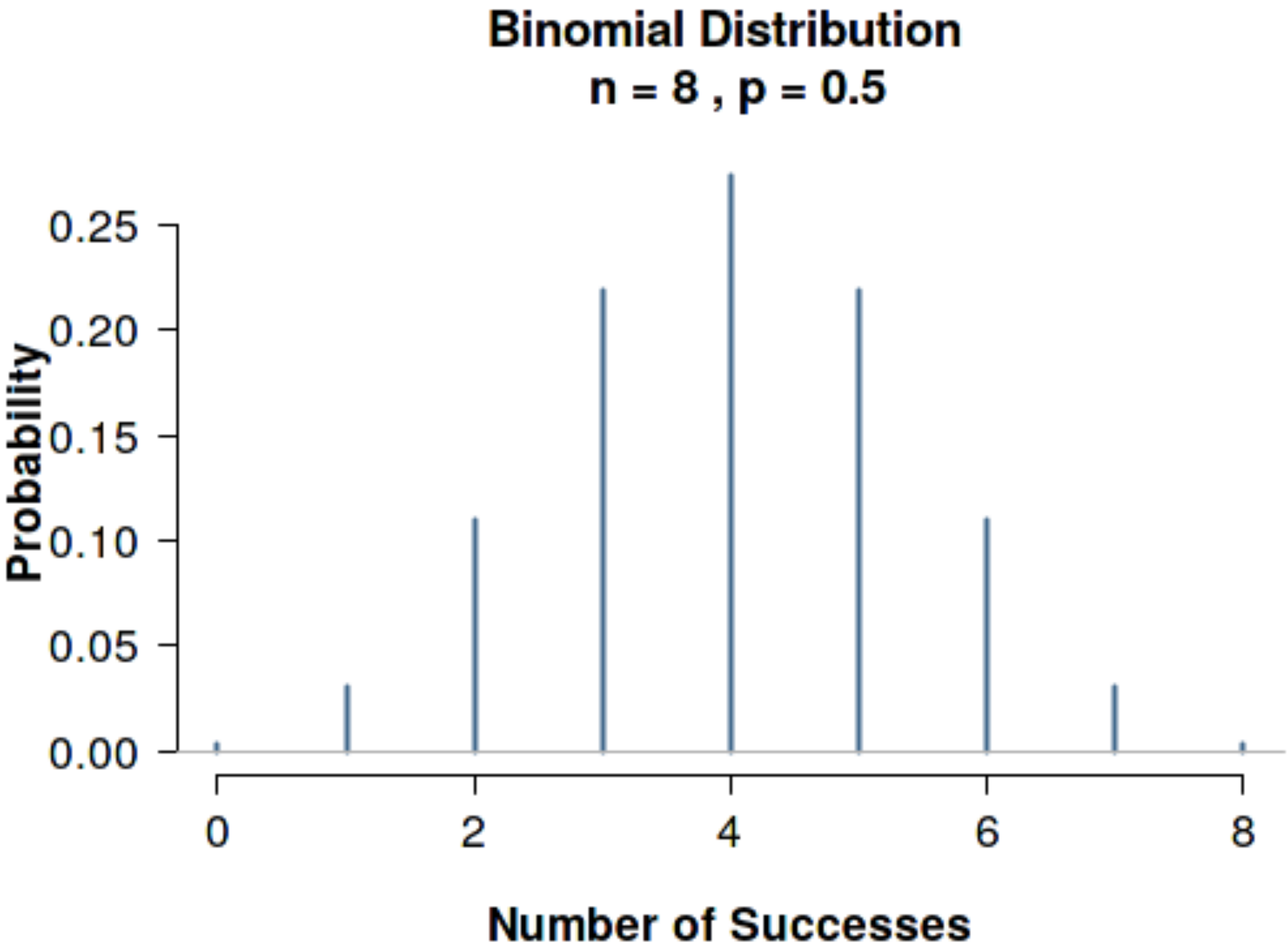


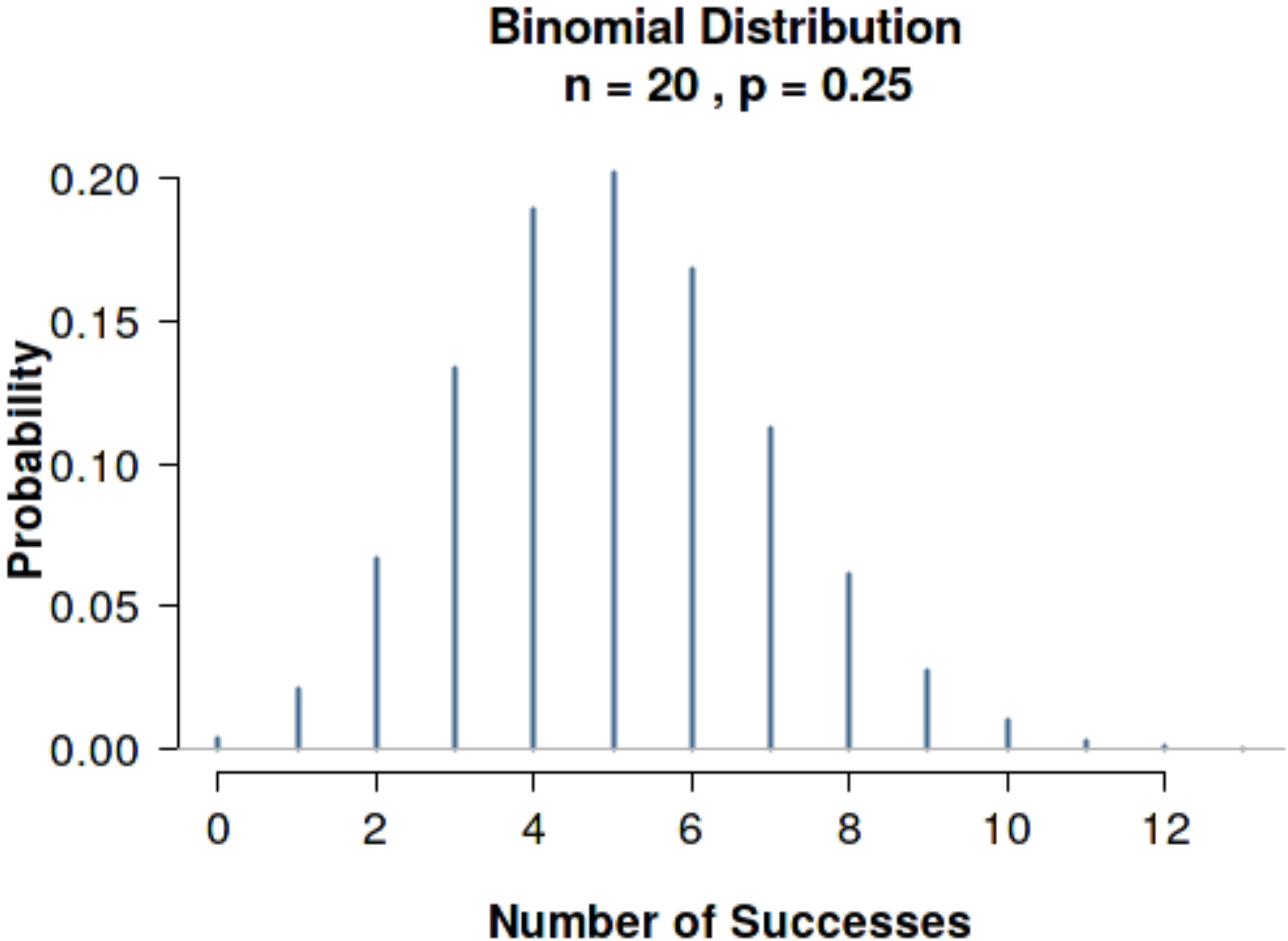
5 Rules for Binomial Distribution

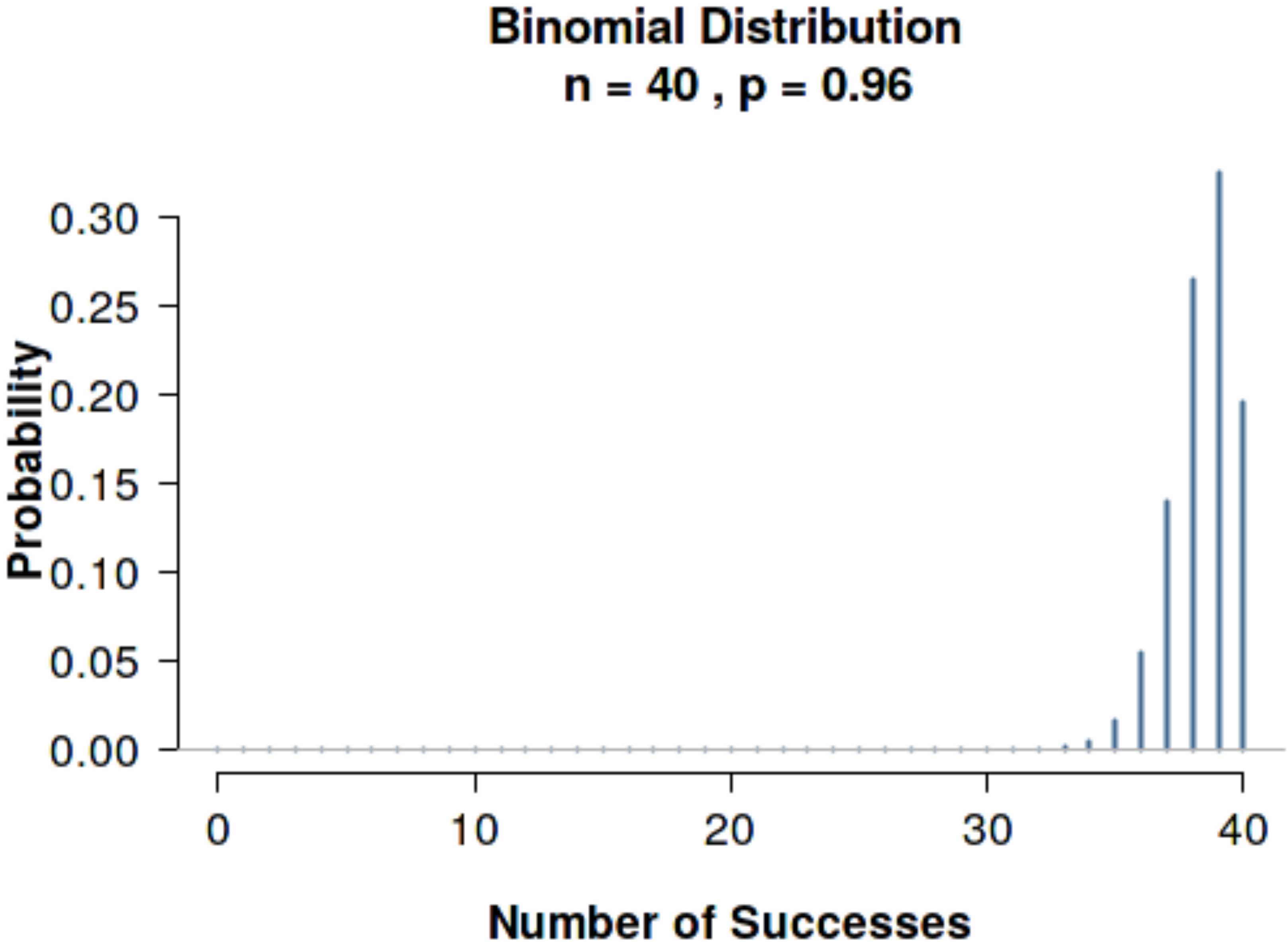


1. Series of *N* trials
2. Only 2 outcomes
3. Outcomes are *mutually exclusive*
4. Outcomes are *independent*
5. Expectation for *P* remains consistent

Does flipping a fair coin meet these?


Binomial Distribution
n = 8 , p = 0.5
Probability
0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
Number of Successes

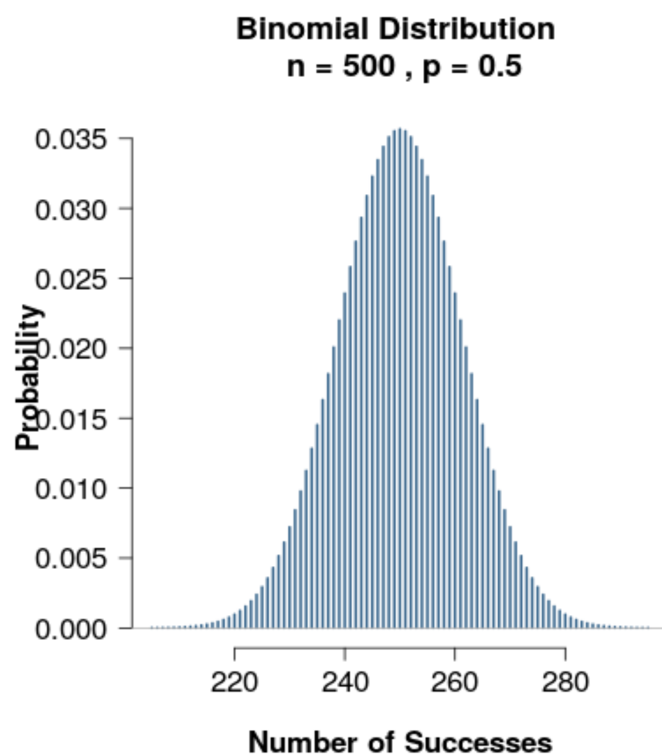

Binomial Distribution
n = 20 , p = 0.25
Probability
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
12
Number of Successes


Binomial Distribution
n = 40 , p = 0.96
Probability
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
Number of Successes

Binomial vs. Normal Distribution

- As *N* increases, binomial distribution *approximates* normal, *z*-distribution

https://shiny.rit.albany.edu/stat/binomial/



Binomial **Table**

skew still represented in tables?

- We don’t need to plot all outcomes, *there’s a **table** for that!*
  - *Appendix D*, Table B (pp. 595-599)

- What is the probability of getting 16 heads out of 19 coin flips? p(16 heads)
  - N = 19
  - P = .50
  - Answer = .0018

table B Binomial distribution—*cont'd*

N	No. of *P* or *Q* Events	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
		P or *Q*									
18	12	.0000	.0000	.0000	.0000	.0002	.0012	.0047	.0145	.0354	.0708
	13	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0045	.0134	.0327
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011	.0039	.0117
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0009	.0031
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
19	0	.3774	.1351	.0456	.0144	.0042	.0011	.0003	.0001	.0000	.0000
	1	.3774	.2852	.1529	.0685	.0268	.0093	.0029	.0008	.0002	.0000
	2	.1787	.2852	.2428	.1540	.0803	.0358	.0138	.0046	.0013	.0003
	3	.0533	.1796	.2428	.2182	.1517	.0869	.0422	.0175	.0062	.0018
	4	.0112	.0798	.1714	.2182	.2023	.1491	.0909	.0467	.0203	.0074
	5	.0018	.0266	.0907	.1636	.2023	.1916	.1468	.0933	.0497	.0222
	6	.0002	.0069	.0374	.0955	.1574	.1916	.1844	.1451	.0949	.0518
	7	.0000	.0014	.0122	.0443	.0974	.1525	.1844	.1797	.1443	.0961
	8	.0000	.0002	.0032	.0166	.0487	.0981	.1489	.1797	.1771	.1442
	9	.0000	.0000	.0007	.0051	.0198	.0514	.0980	.1464	.1771	.1762
	10	.0000	.0000	.0001	.0013	.0066	.0220	.0528	.0976	.1449	.1762
	11	.0000	.0000	.0000	.0003	.0018	.0077	.0233	.0532	.0970	.1442
	12	.0000	.0000	.0000	.0000	.0004	.0022	.0083	.0237	.0529	.0961
	13	.0000	.0000	.0000	.0000	.0001	.0005	.0024	.0085	.0233	.0518
	14	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0082	.0222
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0022	.0074
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

What is p(5 *or more* heads)?

$N = 7, P = .50$

table B Binomial distribution

	No. of P or Q	P or Q									
N	**Events**	**.05**	**.10**	**.15**	**.20**	**.25**	**.30**	**.35**	**.40**	**.45**	**.50**
7	0	.6983	.4783	.3206	.2097	.1335	.0824	.0490	.0280	.0152	.0078
	1	.2573	.3720	.3960	.3670	.3115	.2471	.1848	.1306	.0872	.0547
	2	.0406	.1240	.2097	.2753	.3115	.3177	.2985	.2613	.2140	.1641
	3	.0036	.0230	.0617	.1147	.1730	.2269	.2679	.2903	.2918	.2734
	4	.0002	.0026	.0109	.0287	.0577	.0972	.1442	.1935	.2388	.2734
	5	.0000	.0002	.0012	.0043	.0115	.0250	.0466	.0774	.1172	.1641
	6	.0000	.0000	.0001	.0004	.0013	.0036	.0084	.0172	.0320	.0547
	7	.0000	.0000	.0000	.0000	.0001	.0002	.0006	.0016	.0037	.0078

p(5 or more) = p(5) + p(6) + p(7)

.1641 + .0547 + .0078 = .2266

What is *p*(5 *or more* heads)?

$N = 7, P = .50$

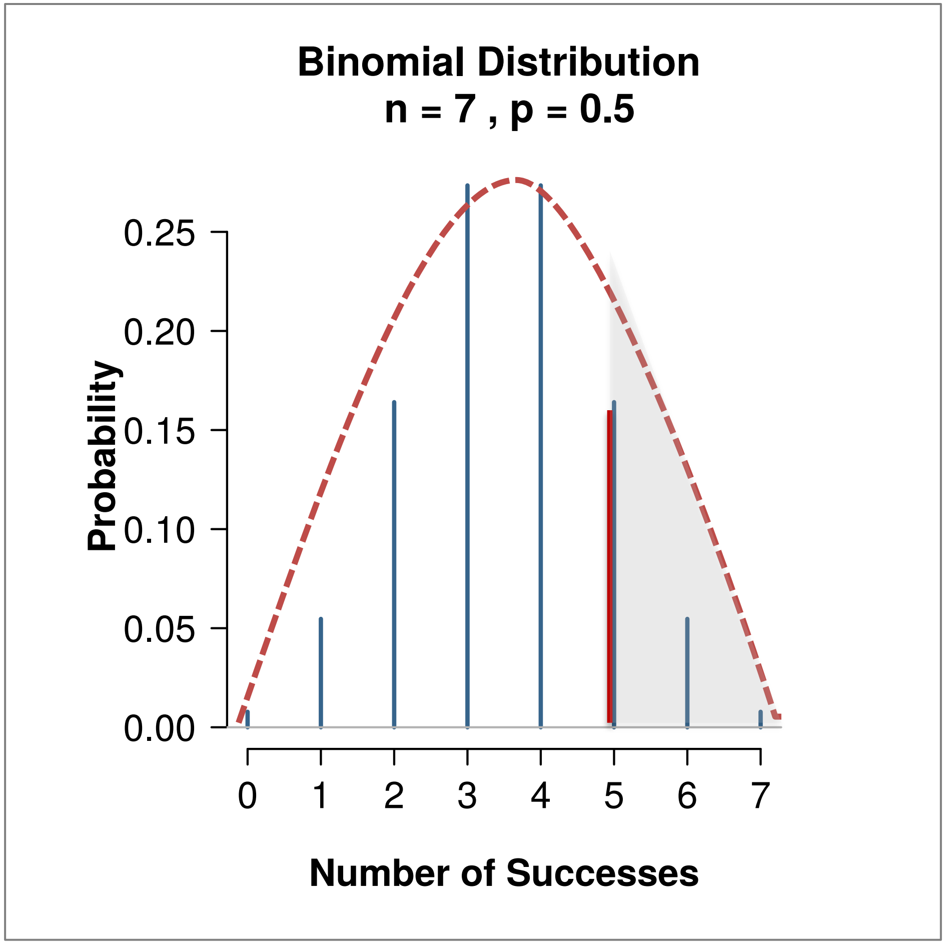
table B Binomial distribution

N	No. of P or Q Events	P or Q .05	.10	.15	.20	.25	.30	.35	.40	.45	.50
7	0	.6983	.4783	.3206	.2097	.1335	.0824	.0490	.0280	.0152	.0078
	1	.2573	.3720	.3960	.3670	.3115	.2471	.1848	.1306	.0872	.0547
	2	.0406	.1240	.2097	.2753	.3115	.3177	.2985	.2613	.2140	.1641
	3	.0036	.0230	.0617	.1147	.1730	.2269	.2679	.2903	.2918	.2734
	4	.0002	.0026	.0109	.0287	.0577	.0972	.1442	.1935	.2388	.2734
	5	.0000	.0002	.0012	.0043	.0115	.0250	.0466	.0774	.1172	.1641
	6	.0000	.0000	.0001	.0004	.0013	.0036	.0084	.0172	.0320	.0547
	7	.0000	.0000	.0000	.0000	.0001	.0002	.0006	.0016	.0037	.0078

$p(5 \text{ or more}) = p(5) + p(6) + p(7)$

$.1641 + .0547 + .0078 = .2266$

What is p(5 *or more* heads)?

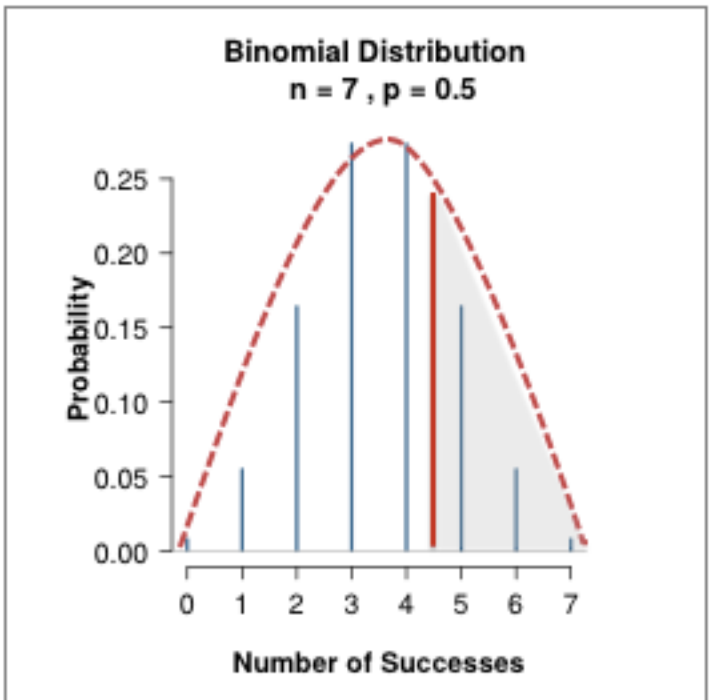


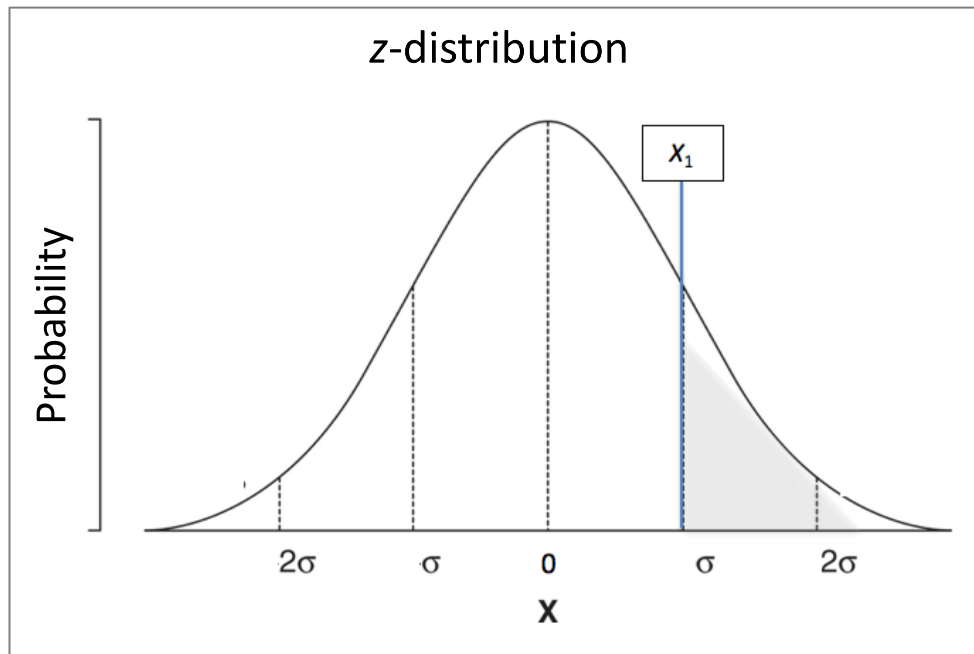
p(5 or more) = .1641 + .0547 + .0078 = .2266

.2266 is the ***area under the curve*** when, $P \geq 5$

Evaluating probability in the tail of the distribution

- What is the probability of 5 heads, or even more?
  Like 6 or 7 heads…or like being taller than 6'…





Inference from Binomial Distribution: Experimental Example (De Obaldia, 2022)

- **Research Question:** Are mosquitoes attracted to certain people?
- **Method:** *Arm-in-Cage* test tournament!
- $N = 20$, $P(\text{bite } X\ )$ = ??

5 Rules for Binomial Distribution

1. Series of *N* trials
2. Only 2 outcomes
3. Outcomes are *mutually exclusive*
4. Outcomes are *independent*
5. Probability of *P* remains consistent

Results

- X_{33} bites = 15
- X_{28} bites = 5
  - $P_{\text{expectation}}(\text{bite } X_{33}) = .50$
  - $P_{observed}(\text{bite } X_{33}) = .75$
- How strange is this result, assuming our expected probability?

p value for this result is .02, so is is most likely that this result is not due to chance

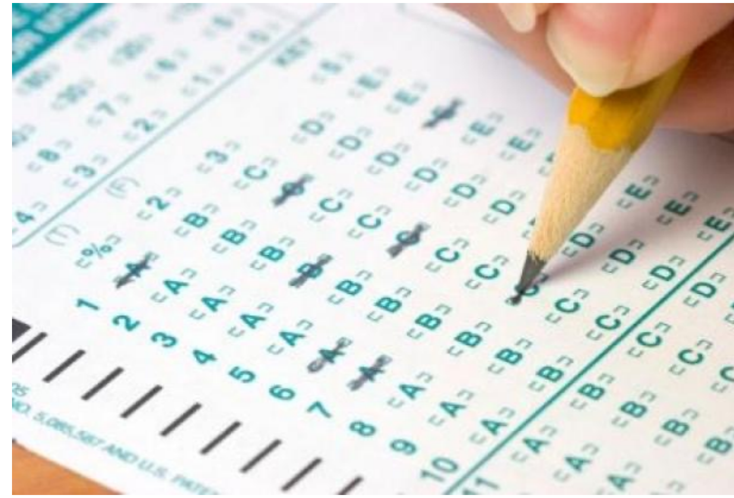
- X_{25} bites = 12
- X_{28} bites = 8

Another example: Multiple choice

[dolphin taking scantron test]

	A	B	C	D	E
1.	[□]	[□]	[□]	[□]	[■]
2.	[□]	[□]	[□]	[□]	[■]
3.	[□]	[□]	[□]	[□]	[■]
4.	[□]	[□]	[□]	[□]	[■]
5.	[□]	[□]	[□]	[□]	[■]

5 Rules

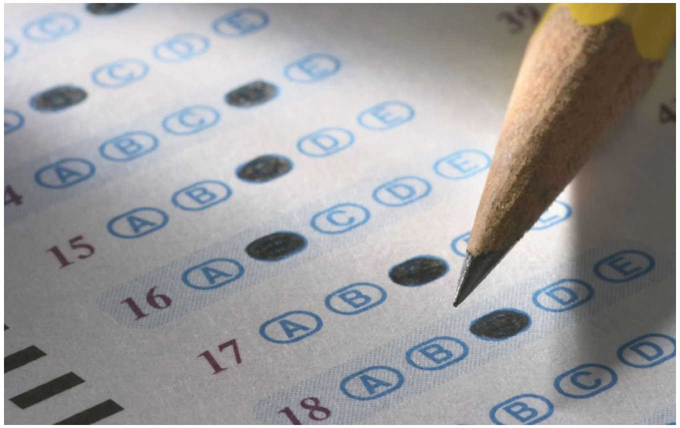


1. Series of *N* trials
2. Only 2 outcomes
3. Outcomes are *mutually exclusive*
4. Outcomes are *independent*
5. Probability of *P* remains consistent

Does a multiple choice test meet these?

Not usually! We'll need to make a few assumptions

Will they pass?



Student needs <u>at least</u> 50% score

$p(>4.5 \text{ correct})$

– $N = 9$ questions

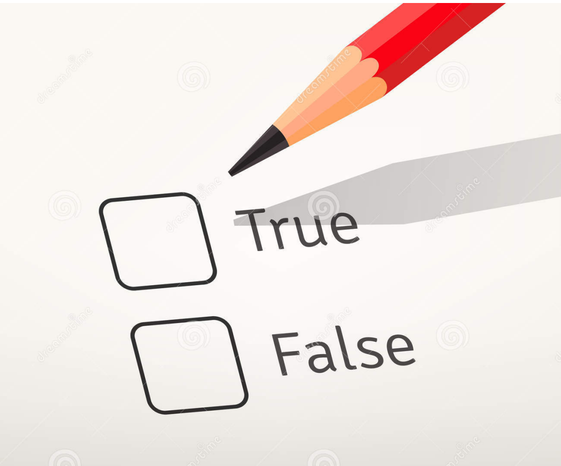
– $P(\text{correct}) = .20$

- Choices a, b, c, d, & e

– Answer: $p(5) + p(6) + p(7) + p(8) + p(9) =$

$.0165 + .0028 + .0003 + .0000 + .0000 = \underline{.0196}$

Will they pass?



- Student needs <u>at least</u> 50% score
  - N = 20 questions
  - P(correct) = .50
  - Answer: $p(10) + \dots p(20)$ = <u>.5881</u>