Hypothesis Testing group activity

Recap: hypothesis testing framework

- We start with a *null hypothesis* (H_0) that represents the status quo.

- We also have an *alternative hypothesis* (H_A) that represents our research question, i.e. what we're testing for.

- We conduct a hypothesis test under the assumption that the null hypothesis is true, either via simulation or traditional methods based on the central limit theorem.

- If the test results suggest that the data do not provide convincing evidence for the alternative hypothesis, we stick with the null hypothesis. If they do, then we reject the null hypothesis in favor of the alternative.

Decision errors (cont.)

There are two competing hypotheses: the null and the alternative. In a hypothesis test, we make a decision about which might be true, but our choice might be incorrect.

		Decision	
		fail to reject H_0	reject H_0
Truth	H_0 true	✓	*Type 1 Error*
	H_A true	*Type 2 Error*	✓

- A *Type 1 Error* is rejecting the null hypothesis when H_0 is true.
- A *Type 2 Error* is failing to reject the null hypothesis when H_A is true.

We (almost) never know if H_0 or H_A is true, but we need to consider all possibilities.

Let's test a new hypothesis

- A student is taking a biology class that studies animal behaviour and is assigned the following research:
- In a certain species, male ducks have green heads and females are all grey. The purpose of the green colouring of the male heads is to attract the females.
- The question is: are female ducks also attracted to the green colour in food, for example in bread?

Let's test a new hypothesis

Idea 1: Female ducks are indifferent to plain versus green bread.

Idea 2: Female ducks prefer green bread.

p: Probability of selecting the green bread first

Let’s test a new hypothesis

Idea 1: Female ducks are indifferent to plain versus green bread.

$p=0.5$

Idea 2: Female ducks prefer green bread.

$p>0.5$

p: Probability of selecting the green bread first

Do ducks prefer green?

The student designs a study in order to be able to make a decision about the two statistical hypotheses.

She will go to a lake near campus where mallards are quite abundant and will randomly select 10 female ducks.

Each duck will be offered two pieces of bread: one plain and one dyed green.

The student will write down which piece of bread each duck approaches first.

Then she will summarize her information reporting how many ducks approach the green bread first.

Pick a test statistic

Think about the variable X = # of ducks in the sample that prefer the green bread. Think of ‘picking green first’ as ‘success’.

Note that the sample size is n=10. If the ducks are truly indifferent to plain versus green bread, what is the distribution of the variable X?

Distribution of the test statistic under the null hypothesis

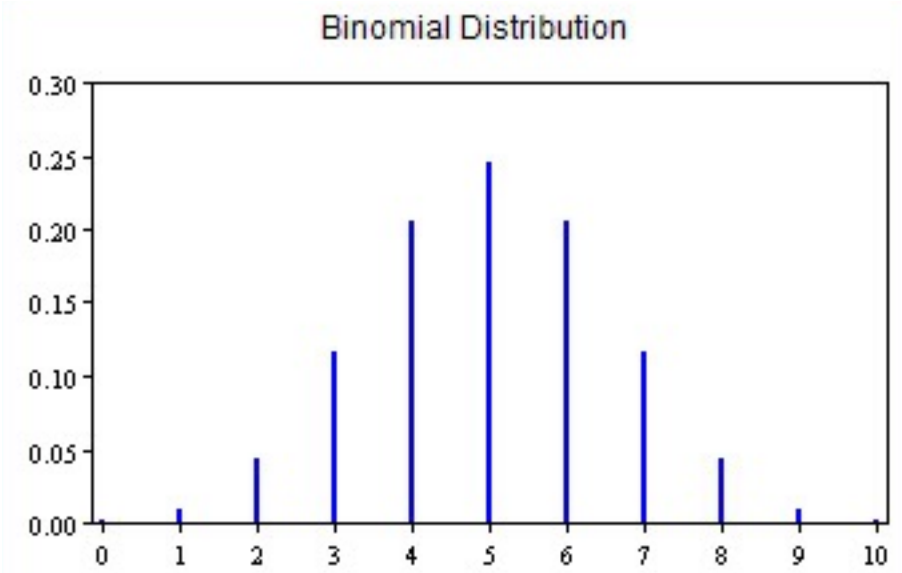
Note that the sample size is n=10. If the ducks are truly indifferent to plain versus green bread, what is the distribution of the variable X?

Distribution of the test statistic under the null hypothesis

Note that the sample size is n=10. If the ducks are truly indifferent to plain versus green bread, what is the distribution of the variable X?

$$X \sim Binom(10, 0.5)$$

x	p(x)
0	0.000977
1	0.009766
2	0.043945
3	0.117188
4	0.205078
5	0.246094
6	0.205078
7	0.117188
8	0.043945
9	0.009766
10	0.000977



Arriving at a conclusion

If female ducks were truly indifferent between green and plain bread, about how many ducks, of the ten that were observed, would you have expected to choose the green bread first?

Suppose you observe that 9 out of 10 ducks selected the green bread.

What is the p-value?

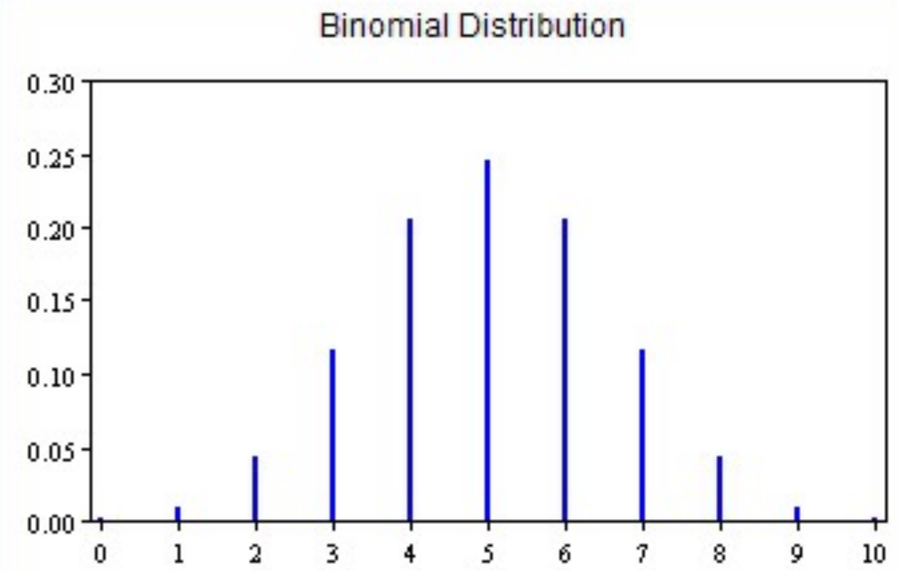
Probability of observing a value at least as unfavorable to the null hypothesis as the one we observed.

Arriving at a conclusion

Suppose you observe that 9 out of 10 ducks selected the green bread.

What is the p-value?

x	p(x)
0	0.000977
1	0.009766
2	0.043945
3	0.117188
4	0.205078
5	0.246094
6	0.205078
7	0.117188
8	0.043945
9	0.009766
10	0.000977



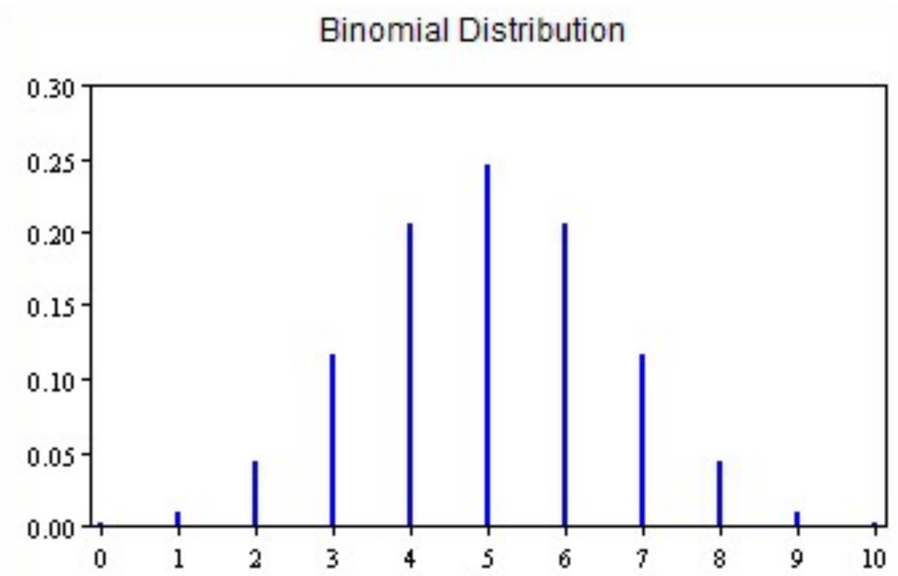
Arriving at a conclusion

Suppose you observe that 9 out of 10 ducks selected the green bread.

What is the p-value?

$$P(X = 9 \; or \; X = 10 | X \sim Binom(10, 0.5))$$

x	p(x)
0	0.000977
1	0.009766
2	0.043945
3	0.117188
4	0.205078
5	0.246094
6	0.205078
7	0.117188
8	0.043945
9	0.009766
10	0.000977



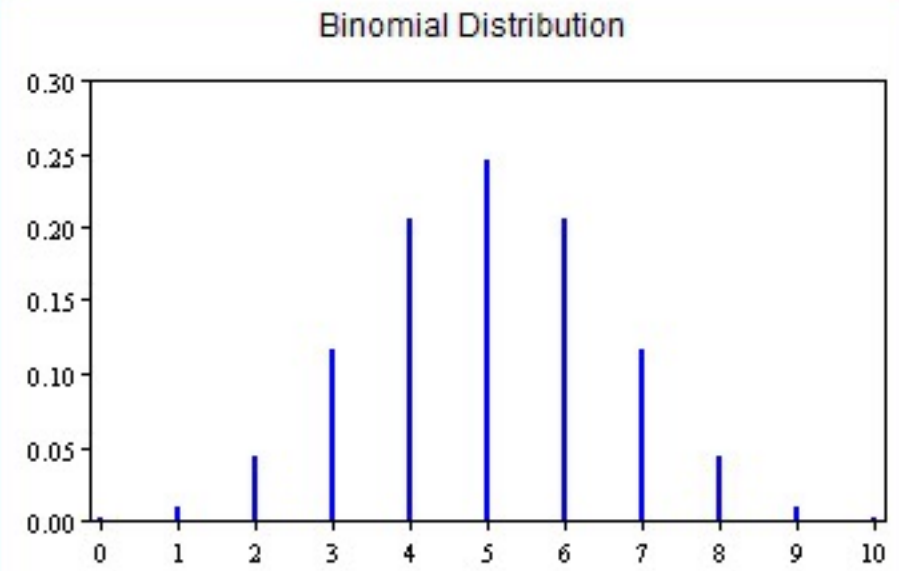
Arriving at a conclusion

Suppose you observe that 9 out of 10 ducks selected the green bread.

What is the p-value?

$$P(X = 9 \; or \; X = 10 | X \sim Binom(10, 0.5)) = 0.011$$

x	p(x)
0	0.000977
1	0.009766
2	0.043945
3	0.117188
4	0.205078
5	0.246094
6	0.205078
7	0.117188
8	0.043945
9	0.009766
10	0.000977



Arriving at a conclusion

To summarize our results we would say that the probability of getting a result as the one we got (9 ducks picking the green bread first) or a more extreme one when the null hypothesis (p=0.5 , meaning ducks are indifferent between green and plain) is true is **0.011** *(This probability of getting the result we got or a more extreme one is called* ***'p-value'****.)*

Should we reject the null hypothesis based on our results?

Possible Errors

Describe in words (in terms of what the ducks prefer and what we say they prefer) each one of these situations?

We select H_a but it is the wrong decision because H_0 is true.

We select H_0 but it is the wrong decision because H_0 is not true.

Effect of the sample size

Students B,C will repeat the experiment with 20 and 40 ducks, respectively

They will use the same $\alpha = 0.05$.

Find, for each of the students, the minimum number of ducks that would have to pick the green bread first in order to reject the null hypothesis.

Fill in the following table where x denotes the number of ducks in the sample that pick the green bread first

Student	*n*	H_0 *will not be rejected if* $x \leq$	H_0 *will be rejected if* $x \geq$
A	10		
B	20		
C	40		

Calculate the probability of a Type II error

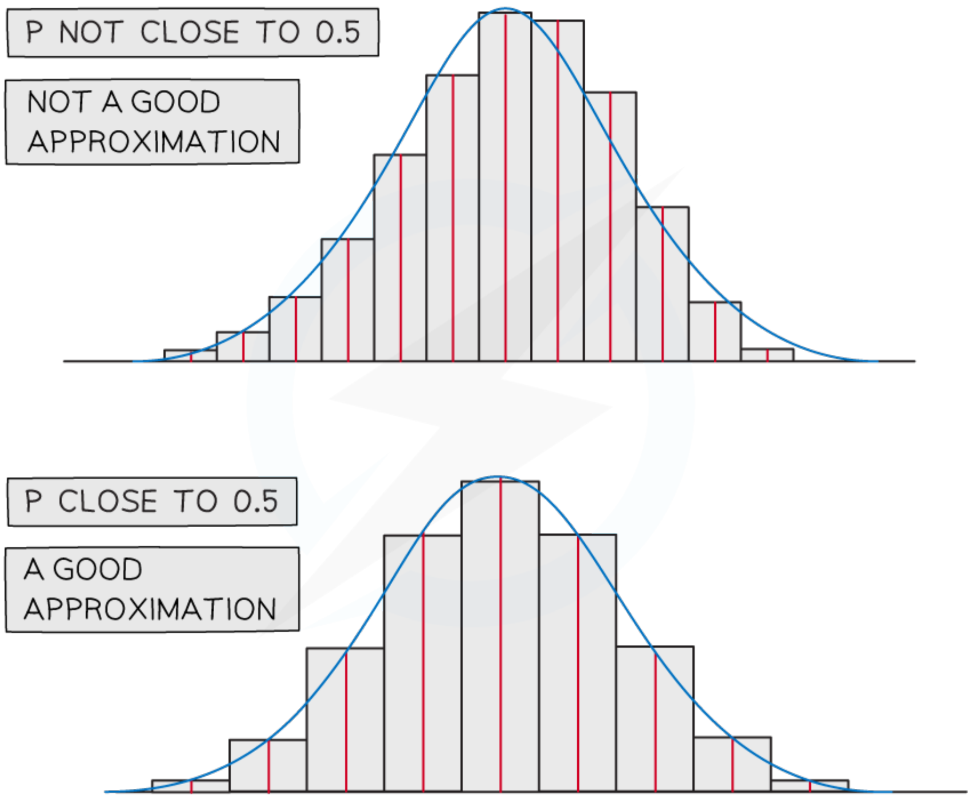
Now assume that H_0 is not true, let us say that 80% of the female mallards prefer the green bread over the plain one.

Taking into consideration the values of x that you wrote in the table in section 4 and using the corresponding for each one of the three sample sizes. Also calculate $1 - \beta = \mathrm{P}($ rejecting $H_0{:}\, p = 0.5$ when $p = 0.8$ $)$

Do a plot of the values of β vs. n. What happens with β as n increases?

Reminder: CLT Application: Binomial

- X_i: Bernoulli with parameter p
- S_n: $X_1 + \cdots + X_n \sim Bin(n, p)$
- CLT: $\frac{S_n - np}{\sqrt{n\,p(1-p)}} \sim N(0,1)$



A binomial distribution $X \sim \mathrm{B}(n, p)$ can be approximated by a normal distribution $X_N \sim \mathrm{N}(\mu, \sigma^2)$ provided

n is large or

p is close to 0.5

Binomial Probabilities

n	k	$p = 0.1$	$p = 0.2$	$p = 0.3$	$p = 0.4$	$p = 0.5$
10	0	.3487	.1074	.0282	.0060	.0010
	1	.3874	.2684	.1211	.0403	.0098
	2	.1937	.3020	.2335	.1209	.0439
	3	.0574	.2013	.2668	.2150	.1172
	4	.0112	.0881	.2001	.2508	.2051
	5	.0015	.0264	.1029	.2007	.2461
	6	.0001	.0055	.0368	.1115	.2051
	7	.0000	.0008	.0090	.0425	.1172
	8	.0000	.0001	.0014	.0106	.0439
	9	.0000	.0000	.0001	.0016	.0098
	10	.0000	.0000	.0000	.0001	.0010

k	$p = 0.1$	$p = 0.2$	$p = 0.3$	$p = 0.4$	$p = 0.5$
0	.1216	.0115	.0008	.0000	.0000
1	.2701	.0576	.0068	.0005	.0000
2	.2852	.1369	.0278	.0031	.0002
3	.1901	.2054	.0716	.0123	.0011
4	.0898	.2182	.1304	.0350	.0046
5	.0319	.1746	.1789	.0746	.0148
6	.0089	.1091	.1916	.1244	.0370
7	.0020	.0545	.1643	.1659	.0739
8	.0003	.0222	.1144	.1797	.1201
9	.0001	.0074	.0654	.1597	.1602
10	.0000	.0020	.0308	.1171	.1762
11	.0000	.0005	.0120	.0710	.1602
12	0000	.0001	.0039	.0355	.1201
13	.0000	.0000	.0010	.0146	.0739
14	.0000	.0000	.0002	.0049	.0370
15	.0000	.0000	.0000	.0013	.0148
16	.0000	.0000	.0000	.0003	.0046
17	.0000	.0000	.0000	.0000	.0011
18	.0000	.0000	.0000	.0000	.0002
19	.0000	.0000	.0000	.0000	.0000
20	.0000	0000	.0000	0000	.0000