Hypothesis Testing (or significance test)

Stastical tests to see if a difference we observe is due to chance.

General steps to perform a hypothesis test:

1 - Formulate your Null and Alternative Hypothesis

- The **H0-null hypothesis** is a hypothesis of no effect. It's the dull boring hypothesis that says that nothing interesting is going on.
- **Ha-Alternative hypothesis** is the opposite of the null. It;s what you're trying to test.

2 - Compare your observed data and expected data and calculate the **test statistic**

3 - Calculate the probability of getting the data you got or something even more extreme if the null were true. This called the **p-value**.

4 - Make your conclusion and interpret it in the context of the problem. If p is very low, we say that the data support rejecting the null hypothesis.

The One Sample Z Test: One-sided Hypothesis

The first type of hypothesis test we are going to look at is the one-sample z-test. You can do a z-test for means or for proportions. This is the most simple type of hypothesis test and it uses z-scores and the normal curve. Let's look at one below!

Hypothesis Test Example: Suppose a large university claims that the average ACT score of their incoming freshman class is **30**, but we think the University may be inflating their average. To test the University's claim we take a simple random sample of **50 students** and find their **average to be only 28.3** with an **SD of 4**. Perform a hypothesis test to test the claim. Here are the 4 steps:

1. Formulate your Null and Alternative Hypotheses.
    - **Ho- Null Hypothesis:** The true average ACT score of all freshman is 30 as claimed.
        - can be written in symbols as well: Ho: μ = 30
        - μ is the symbol for the population mean
    - **Ha- Alternative Hypothesis:** The true average ACT score of all freshman is less than 30.
        - This can be written in symbols as well: Ha: μ < 30
2. Our **test statistic** for the one sample z test is z! We can calculate z using our z-score formula for random variables since we are dealing with a sample of 50 students.

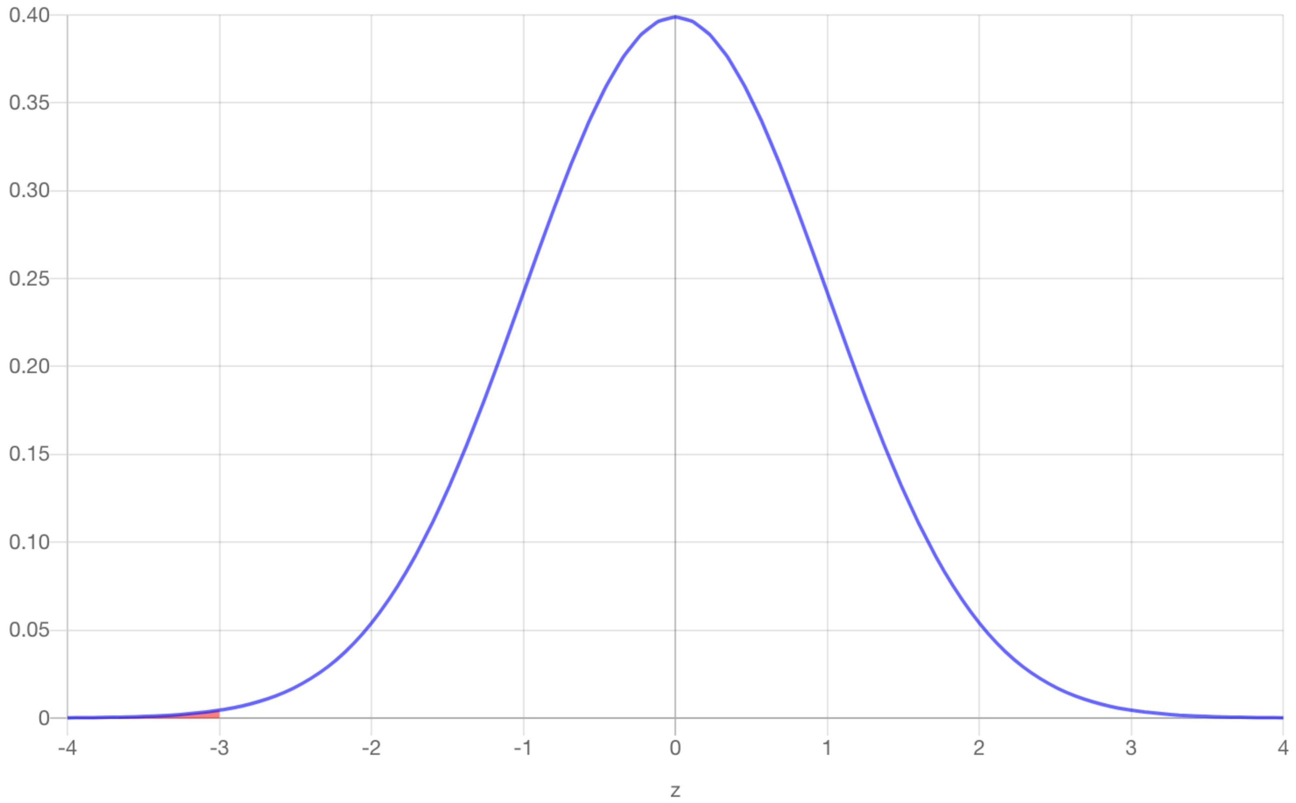
$$Z = \frac{\text{value - EV}}{\text{SE}}$$

- In our case, the expected value (EV) is 30 since we are assuming our null hypothesis is true (until proven otherwise).
- Since we are dealing with means, our SE is found using the following formula:

Our z-score is -3. See the calculation below:

$$Z = \frac{28.3 - 30}{\frac{4}{\sqrt{50}}} = -3$$

3. Calculate the probability of getting the data you got or something even more extreme if the null were true. This is called the **p-value**. In this case, our p-value is going to be the area to the left of z = -3. We can use Python to calculate this by using **norm.cdf(-3).**
    - We get that the p-value is 0.0013.
    - This is the probability that we would get a sample average of 28.3 given that the null hypothesis was true (the true average was 30).



4. Make your conclusion and interpret it in the context of the problem. If p is very low, we say that the data support rejecting the null hypothesis.
   - Our p-value is less than 5% so we reject our Null Hypothesis. In other words, there is evidence of the Alternative Hypothesis (that the University is inflating their average).

The One Sample Z Test: Two-sided Hypothesis

Hypothesis Test Example: Now we're going to test the above claim but with a different alternative hypothesis. The large university still claims that the average ACT score of their incoming freshman class is 30, but now we think the University may be inflating **or** deflating their average. To test the University's claim we take a simple random sample of 50 students and find their average to be only 28.3 with an SD of 4. Perform a hypothesis test to test the claim with our new alternative hypothesis. Here are the 4 steps:

1. Formulate your Null and Alternative Hypotheses.

- **Ho- Null Hypothesis**: The true average ACT score of all freshman is 30 as claimed.
  - This can be written in symbols as well: Ho: μ = 30
  - μ is the symbol for the population mean
- **Ha- Alternative Hypothesis**: The true average ACT score of all freshman is less than 30 **or** greater than 30.
  - This can be written in symbols as well: Ho: μ != 30

2. Step 2 is the same as the one-sided example, so our z score is still -3.
3. Calculate the probability of getting the data you got or something even more extreme if the null were true. This is called the **p-value**. In this case, our p-value is going to be the area to the left **or** right of z = -3. We can use Python to calculate this by using **2*norm.cdf(-3)**.
   - We get that the p-value is 0.0027.
   - This is the probability that we would get a sample average of 28.3 given that the null hypothesis was true (the true average was 30).
4. Make your **conclusion** and interpret it in the context of the problem. If p is very low, we say that the data support rejecting the null hypothesis.
   - Our p-value is less than 5% so we reject our Null Hypothesis. In other words, there is evidence of the Alternative Hypothesis (that the University is inflating or deflating their average).

Python Implementation

Statsmodel library installation

- conda => conda install -c conda-forge statsmodels
- pip => python -m pip install statsmodels

```
# import library
from statsmodels.stats.weightstats import ztest
import random
import pandas as pd
import numpy as np
```

statsmodels.stats.weightstats.ztest(x1, x2=None, value=0 alternative='two-sided',usevar='pooled', ddof=1.0)

Test for mean based on the normal distribution, one or two samples In the case of two samples, the samples are assumed to be independent.

Returns:

- tstat - float (test statistic)
- pvalue - float (pvalue of the z-test)

Example 1:

Simulate 100 rolls of an unfair die, that is 3x more likely to roll a 6 than any other roll:

```
data = []
for i in range(100):
  roll = random.choice([1,2,3,4,5,6,6])
  d = {'roll':roll}
  data.append(d)
df = pd.DataFrame(data)
```

```
df.head()
```

	roll
0	6
1	1
2	6
3	3
4	1

Next steps: Generate code with df View recommended plots New interactive sheet

```
df.shape
```

```
(100, 1)
```

Example 2

Use ztest to find if our dice rolls were likely to be from a fair die?

Null Hypothesis (H_0):

- The probability of rolling a 6 is equal to 1/6 (p = 1/6)
- In other words, the die is fair

Alternative Hypothesis (H_1):

- The probability of rolling a 6 is not equal to 1/6 (p ≠ 1/6)
- In other words, the die is unfair

```
mean_val = df['roll'].mean()
# print(mean_val)

t_stat, p_val  = ztest(df['roll'],value=mean_val)
print(t_stat, p_val)
```

```
0.0 1.0
```

```
alpha = 0.5
if p_val < alpha:
  print('Reject Null Hypothesis')
else:
  print('Fail to reject Null Hypothesis')
```

```
Fail to reject Null Hypothesis
```

TRY yourself!

Now try rolling the die 10000 times and test the hypothesis again.

```
Start coding or generate with AI.
```