04 Proportion Testing

November 20, 2022

```
[1]: import pandas as pd
     import numpy as np
     import matplotlib.pyplot as plt
     from scipy import stats
```

1 Bernoulli trial and proportion testing

$$P(S_n = k) = \binom{n}{k} * p^k * (1 - p)^{n-k}, \quad \text{for } 0 \le k \le n$$

1.1 Hypothesis

$$
\begin{aligned}
H_0 : & \quad p_0 = p \\
H_1 : & \quad p > p_0
\end{aligned}
$$

1.2 Test statistics

Assume the following things are given:

- k: E. g. number of failed products
- n: Total number of observations
- p_0: Assumed probability that a product failes during production

$$T = \frac{\bar{x} - p}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

where $\bar{x} = \frac{k}{n}$

```
[2]: k = 46
     n = 2000
     p0 = 0.02
     alpha = 0.01

     x_bar = k / n
     s_x = p0 * (1 - p0)

     # Calculate test statistic
     t = (x_bar - p0) / (np.sqrt(s_x / n))
```

```python
# Calculate rejection area - right-sided
c = stats.norm().ppf(1 - alpha)

# Calculate the p-value
p_value = 1 - stats.norm().cdf(t)

print(f"Test statistic: {t}")
print(f"Rejection area: {c}")
print(f"P-Value:        {p_value}")
print(f"Reject H0:      {t > c}")
```

```
Test statistic: 0.9583148474999096
Rejection area: 2.3263478740408408
P-Value:        0.1689520095515027
Reject H0:      False
```

[ ]: