Statistical Significance

Probability and Statistics for Data Science

Carlos Fernandez-Granda

These slides are based on the book Probability and Statistics for Data Science by Carlos Fernandez-Granda, available for purchase here. A free preprint, videos, code, slides and solutions to exercises are available at https://www.ps4ds.net

Hypothesis testing

1. Choose a conjecture
2. Choose null hypothesis
3. Choose test statistic
4. Decide significance level α
5. Gather data and compute test statistic
6. Compute p value
7. Reject the null hypothesis if p value $\leq \alpha$

Antetokounmpo's free throws

Conjecture: Free throw percentage is higher at home than away

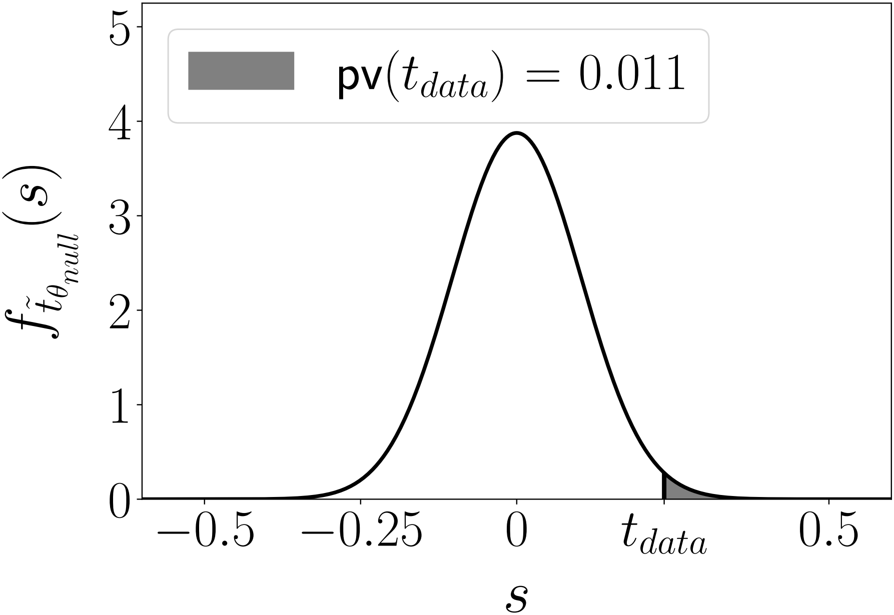
Null hypothesis: Percentage is the same

Test statistic:

$$\frac{\text{Made at home}}{\text{Attempted at home}} - \frac{\text{Made away}}{\text{Attempted away}}$$

Significance level: $\alpha := 0.05$

We reject the null hypothesis



What can go wrong?

Type 1 error: False positive

Null hypothesis holds, but we reject it

Type 2 error: False negative

Null hypothesis does not hold, but we do not reject it

$\mathrm{P}\,(\text{False positive}) \leq \alpha$

1. Choose a conjecture
2. Choose null hypothesis
3. Choose test statistic
4. Decide significance level α
5. Gather data and compute test statistic
6. Compute p value
7. Reject the null hypothesis if p value $\leq \alpha$

False positive

A false positive happens if

1. Null hypothesis holds
2. P value $\leq \alpha$

Parametric testing

Distribution of test statistic depends on parameter θ

Simple null hypothesis: $\theta = \theta_{\text{null}}$

Test statistic under null hypothesis: $\tilde{t}_{\text{null}}$

P value under null hypothesis: $\tilde{u} := \text{pv}(\tilde{t}_{\text{null}})$

$$\mathrm{P}\left(\text{False positive}\right) = \mathrm{P}\left(\tilde{u} \leq \alpha\right)$$

P value under null hypothesis

P-value function

$$\begin{aligned} \mathrm{pv}(t) &:= \mathrm{P}\left(\tilde{t}_{\text{null}} \geq t\right) \\ &= 1 - F_{\tilde{t}_{\text{null}}}(t) \end{aligned}$$

$$\begin{aligned} \tilde{u} &:= \mathrm{pv}(\tilde{t}_{\text{null}}) \\ &= 1 - F_{\tilde{t}_{\text{null}}}(\tilde{t}_{\text{null}}) \end{aligned}$$

What happens if we plug a random variable into its own cdf?

$\tilde{b} := F_{\tilde{a}}(\tilde{a})$

We assume that $F_{\tilde{a}}$ is invertible

For $0 \leq b \leq 1$

$$\begin{aligned} F_{\tilde{b}}(b) &= \mathrm{P}\left(\tilde{b} \leq b\right) \\ &= \mathrm{P}\left(F_{\tilde{a}}(\tilde{a}) \leq b\right) \\ &= \mathrm{P}\left(\tilde{a} \leq F_{\tilde{a}}^{-1}(b)\right) \\ &= F_{\tilde{a}}\left(F_{\tilde{a}}^{-1}(b)\right) \\ &= b \end{aligned}$$

Uniform distribution in $[0, 1]$

$\tilde{c} := 1 - F_{\tilde{a}}(\tilde{a}) = 1 - \tilde{b}$

For $0 \leq c \leq 1$

$$
\begin{aligned}
F_{\tilde{c}}(c) &= \mathrm{P}(\tilde{c} \leq c) \\
&= \mathrm{P}(1 - \tilde{b} \leq c) \\
&= \mathrm{P}(\tilde{b} \geq 1 - c) \\
&= 1 - (1 - c) \\
&= c
\end{aligned}
$$

Uniform distribution in $[0, 1]$

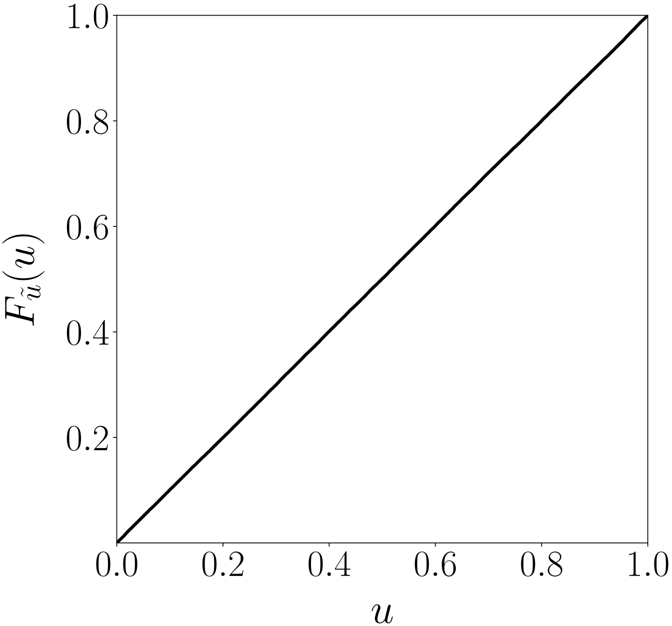
P value under null hypothesis

If $\tilde{t}_{\text{null}}$ is continuous

$$\begin{aligned} \tilde{u} &:= \text{pv}(\tilde{t}_{\text{null}}) \\ &= 1 - F_{\tilde{t}_{\text{null}}}(\tilde{t}_{\text{null}}) \end{aligned}$$

Uniform distribution in $[0, 1]$!

P value under null hypothesis



False positives

$$\tilde{u} := \mathsf{pv}(\tilde{t}_{\mathsf{null}})$$

Uniform distribution in $[0, 1]$

$$\begin{aligned} \mathrm{P}\left(\text{False positive}\right) &= \mathrm{P}\left(\tilde{u} \leq \alpha\right) \\ &= \alpha \end{aligned}$$

Die rolls

Conjecture: Probability of rolling $3 > 1/6$

Null hypothesis: Probability of rolling $3 = 1/6$

Test statistic: Number of 3s (out of 100)

Distribution under null hypothesis?

Binomial with parameters $n := 100$ and $\theta = 1/6$

Significance level: $\alpha := 0.05$

P value: 0.15 $\quad > \alpha$

We don't reject the null hypothesis

Discrete test statistic

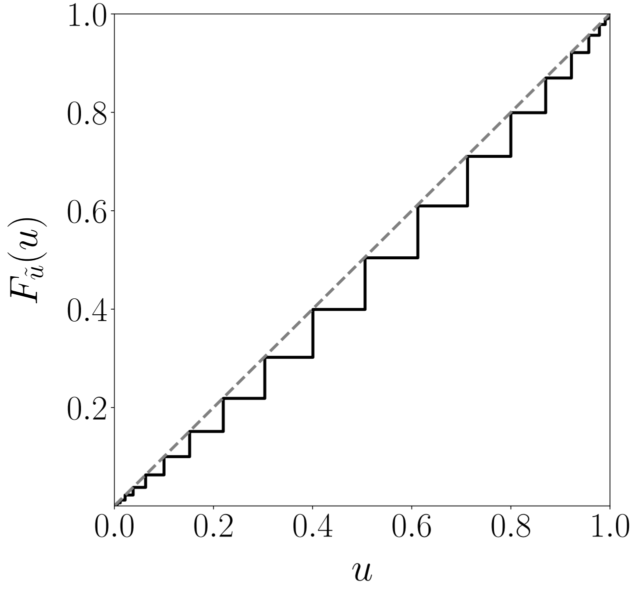
$$\tilde{u} := \mathrm{pv}(\tilde{t}_{\text{null}})$$

For any $u \in [0, 1]$

$$F_{\tilde{u}}(u) \leq u$$

$$\begin{aligned} \mathrm{P}\left(\text{False positive}\right) &= \mathrm{P}\left(\tilde{u} \leq \alpha\right) \\ &= F_{\tilde{u}}(\alpha) \leq \alpha \end{aligned}$$

P value under null hypothesis



Composite null hypothesis

A false positive happens if

1. Null hypothesis holds, $\theta = \theta_0$ for some $\theta_0 \in \Theta_{\text{null}}$
2. P value $\leq \alpha$

$$\begin{aligned}
\mathrm{P}\left(\text{False positive}\right) &= \mathrm{P}\left(\text{pv}(\tilde{t}_{\theta_0}) \leq \alpha\right) \\
&= \mathrm{P}\left(\sup_{\theta \in \Theta_{\text{null}}} \text{pv}_{\theta}(\tilde{t}_{\theta_0}) \leq \alpha\right) \\
&\leq \mathrm{P}\left(\text{pv}_{\theta_0}(\tilde{t}_{\theta_0}) \leq \alpha\right) \\
&\leq \alpha
\end{aligned}$$

What have we learned

P value has a uniform distribution under null hypothesis

Thresholding p value controls the probability of a false positive