The Null Hypothesis and the P Value

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

Plan

Define key concepts in hypothesis testing

- Null hypothesis
- Test statistic
- P value

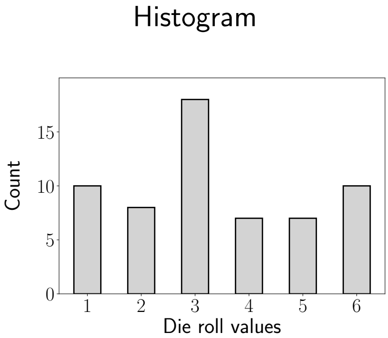
Hypothesis testing

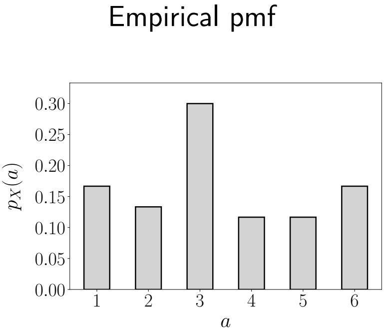
Goal: Determine whether data supports a conjecture

Play devil's advocate: *Maybe it's just chance*

If this is very unlikely $\implies$ Data supports conjecture

60 die rolls





Conjecture? $\mathrm{P}\,(\text{rolling a 3}) > 1/6$

Null hypothesis

Contradicts conjecture

Null hypothesis for die example?

$\mathrm{P}\,(\text{rolling a 3}) = 1/6 \quad \text{or} \quad \mathrm{P}\,(\text{rolling a 3}) \leq 1/6$

Original conjecture is the alternative hypothesis

Test statistic

Function of the data

Large value is evidence against null hypothesis

Test statistic for die example?

Number of 3s

Die example

Test statistic: 18 3s (out of 60)

Is this unlikely under null hypothesis: $\mathrm{P}\left(\text{rolling a 3}\right) = 1/6$?

Yes

End of story? No!

We have selected hypothesis after looking at the data

License plate

Your friend says:

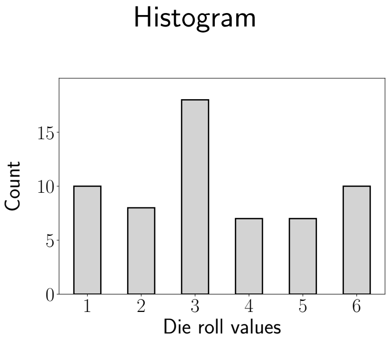
Look, a car with license plate number EMC6055!

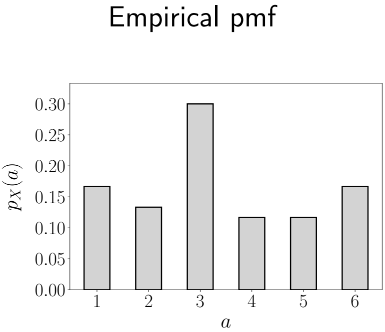
One million cars drive through Manhattan every day

Isn't this amazing? A one-in-a-million chance!

No!

60 die rolls


Histogram
Count
Die roll values
0
5
10
15
1
2
3
4
5
6


Empirical pmf
$p_X(a)$
a
0.00
0.05
0.10
0.15
0.20
0.25
0.30
1
2
3
4
5
6

Hypothesis testing

1. Decide null hypothesis (without looking at the data!)
2. Select test statistic (without looking at the data!)
3. Evaluate test statistic on data
4. Determine how likely it is under null hypothesis

This requires modeling the distribution of the test statistic

Parametric testing

Distribution depends on a small number of parameters θ

Simple null hypothesis: Parameters equal single value $\theta = \theta_{\mathsf{null}}$

Composite null hypothesis: Parameters belong to a set $\theta \in \Theta_{\mathsf{null}}$

Die rolls

Assumption: Rolls are independent

Probability of rolling $3 = \theta$

Test statistic: Number of 3s out of 100 new rolls

Distribution? Binomial with parameters $n := 100$ and θ

Simple null hypothesis: $\theta = \theta_{\text{null}} = \frac{1}{6}$

Composite null hypothesis: $\theta \in \Theta_{\text{null}} = \left[0, \frac{1}{6}\right]$

Observed test statistic: 21

Key question:

How likely is observing 21 3s or more under null hypothesis?

P value

Probability of observing larger or equal test statistic under null hypothesis

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

P-value function

$$\text{pv}(t) := \mathrm{P}\left(\tilde{t}_{\theta_{\text{null}}} \geq t\right)$$

P value of data is $\text{pv}(t_{\text{data}})$

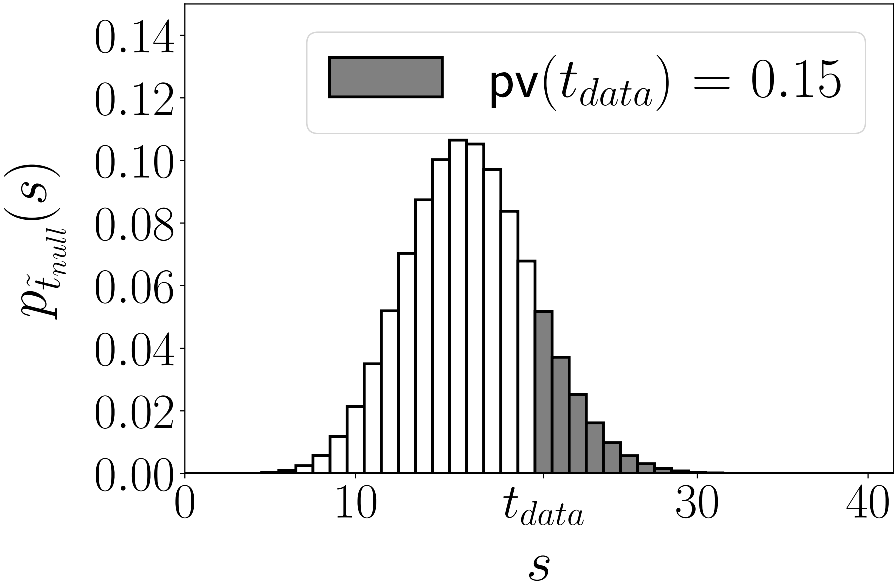
Die rolls

P-value function for $t = 0, 1, \ldots, 100$

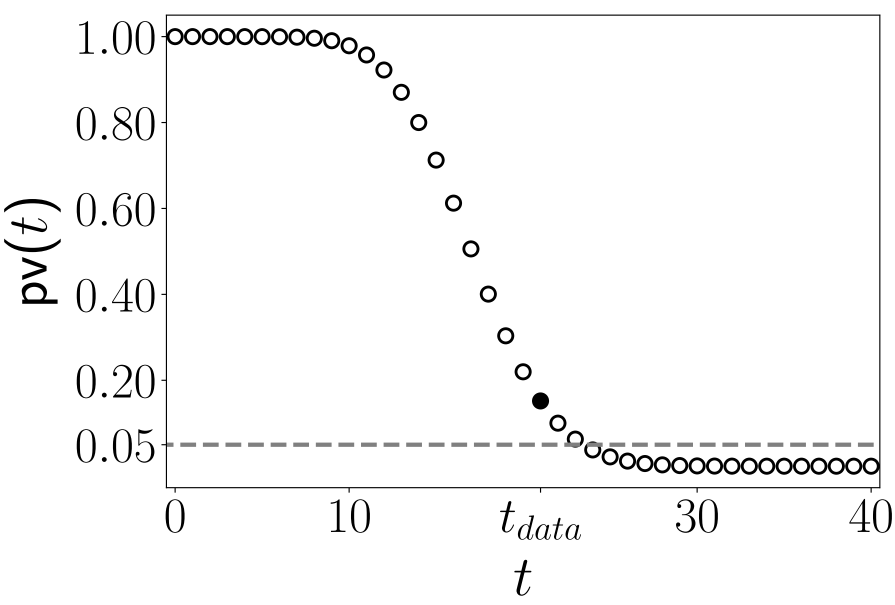
$$\begin{aligned} \mathrm{pv}(t) &:= \mathrm{P}\left(\tilde{t}_{\theta_{\text{null}}} \geq t\right) \\ &= \sum_{i=t}^{100} \binom{100}{i} \left(\frac{1}{6}\right)^{i} \left(\frac{5}{6}\right)^{100-i} \end{aligned}$$

$$\begin{aligned} \mathrm{pv}(t_{\text{data}}) &= \sum_{i=21}^{100} \binom{100}{i} \left(\frac{1}{6}\right)^{i} \left(\frac{5}{6}\right)^{100-i} \\ &= 0.15 \end{aligned}$$

Test statistic under null hypothesis



P-value function



Composite null hypothesis $\theta \in \Theta_{\text{null}}$

P-value function

$$\text{pv}(t) := \sup_{\theta \in \Theta_{\text{null}}} \mathrm{P}\left(\tilde{t}_\theta \geq t\right)$$

P value of data is $\text{pv}(t_{\text{data}})$

Die rolls

P-value function for composite null hypothesis

$$\begin{aligned}
\text{pv}(t) &:= \sup_{0 \leq \theta \leq 1/6} \mathrm{P}\left(\tilde{t}_{\theta} \geq t\right) \\
&= \sup_{0 \leq \theta \leq 1/6} \quad \sum_{i=t}^{100} \binom{100}{i} \theta^{i} \left(1-\theta\right)^{100-i} \\
&= \sum_{i=t}^{100} \binom{100}{i} \left(\frac{1}{6}\right)^{i} \left(\frac{5}{6}\right)^{100-i}
\end{aligned}$$

$$\begin{aligned}
\text{pv}(t_{\text{data}}) &= \sum_{i=21}^{100} \binom{100}{i} \left(\frac{1}{6}\right)^{i} \left(\frac{5}{6}\right)^{100-i} \\
&= 0.15
\end{aligned}$$

P value

Tempting interpretation:

The probability that the die is fair ($\theta = 1/6$) *is* 0.15

Doesn't make sense: θ is deterministic!

P value

Correct interpretation:

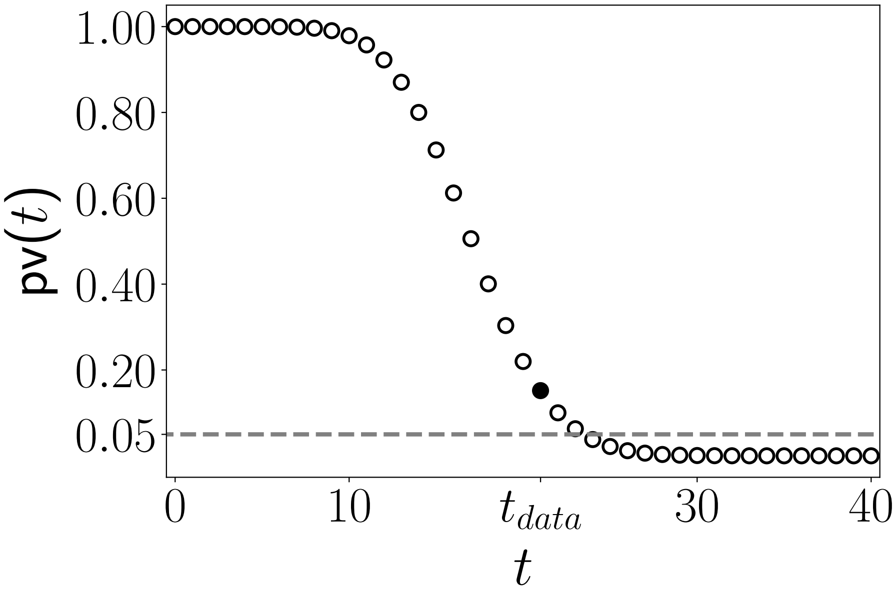
If $\theta = 1/6$, *the probability of observing* $\geq$ *21 3s is* 0.15

Statistical significance

We reject the null hypothesis when p value $\leq \alpha$

Guarantees that probability of a false positive $\leq \alpha$

Die rolls



Die rolls

$\alpha := 0.05 < 0.15$

We don't reject the null hypothesis

Does that mean we think the die is fair? No!

We just don't have enough evidence to rule it out

What have we learned

Key concepts in hypothesis testing

- Null hypothesis
- Test statistic
- P value