P-values
False Discovery Rate
Family-wise Error Rate

Robert Brown

Thank you!

UCLA Computational Medicine

Your test statistic

Write down a number between 1 and 100!

My number: 42

Your number is your "test statistic" it is a draw from the "Give me a number between 1 and 100!" null distribution.

$$S^O = 42$$

How "extreme" is your number?

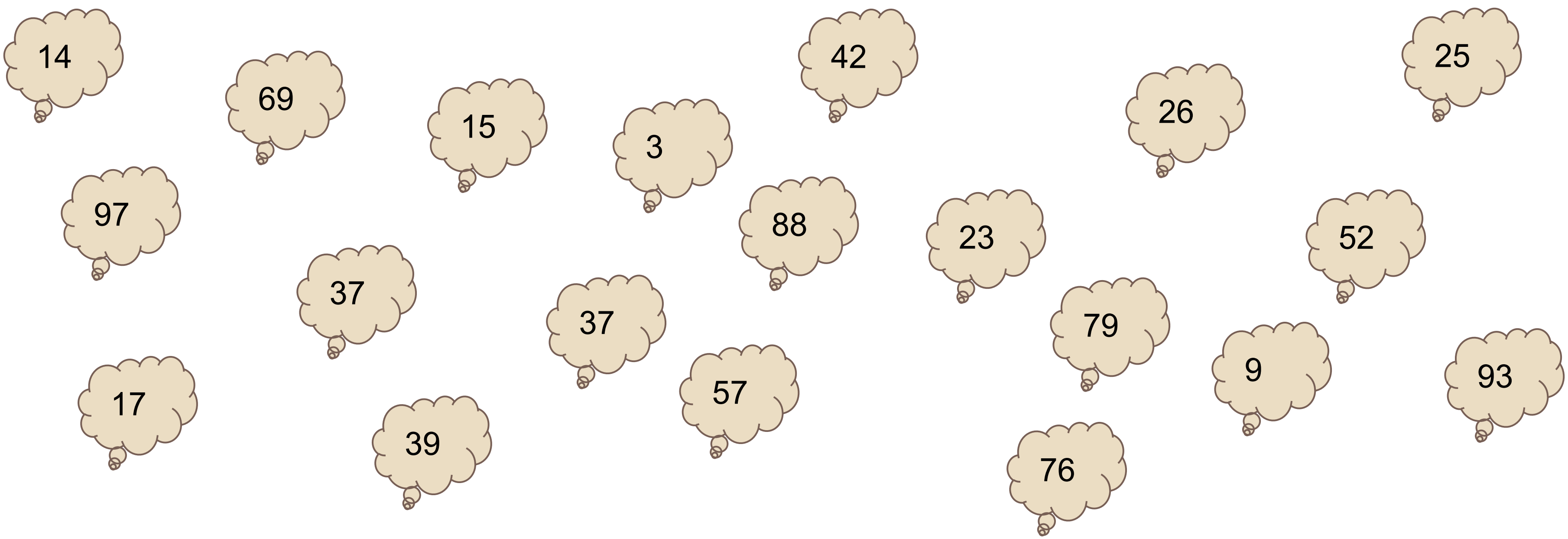
$$P(S^n \text{ more extreme than } S^O = 42 | \text{ 1 to 100 dist})$$

1. Larger is more extreme.
2. Smaller is more extreme.
3. Greater absolute distance from zero is more extreme.
4. More or less than expected is extreme.

Null Distribution (empirical)

Give me a number between 1 and 100!

$$P(S^n \text{ more extreme than } S^O = 42 | \text{ 1 to 100 dist})$$



$$P(S^n \geq S^O = 42 | \text{ 1 to 100 dist})$$

P-value!!!!!!!! (empirical)

Give me a number between 1 and 100!

3 9 14 15 17 23 25 26 37 37

39 42 52 57 69 76 79 88 93 97

42

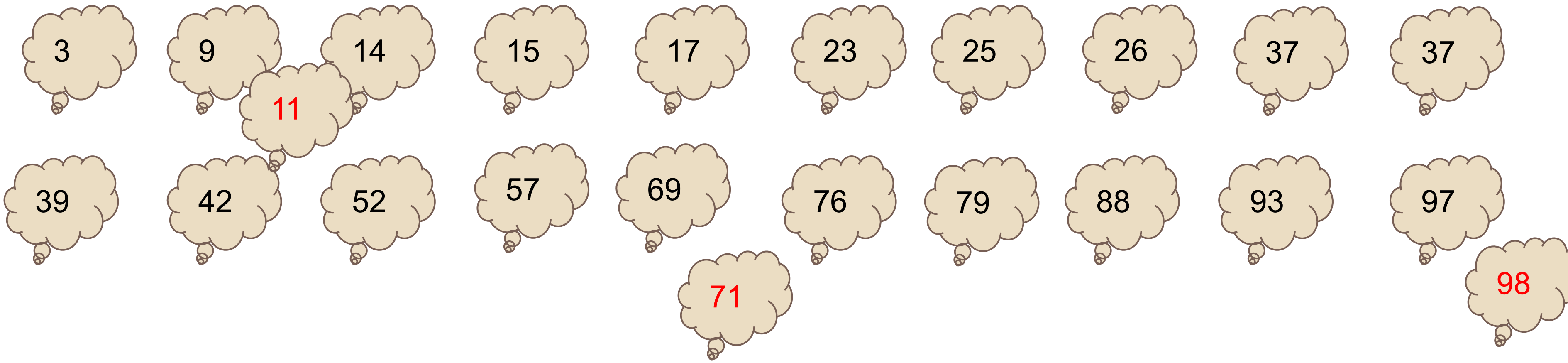
$$P(S^n \geq S^O = 42 |\ 1 \text{ to } 100 \text{ dist})$$

$$P(S^n \geq S^O = 42 |\ 1 \text{ to } 100 \text{ dist}) \approx 9/20 = 0.45$$

$$\text{P-value} = P(S^n \geq S^O = 42 |\ 1 \text{ to } 100 \text{ dist}) \approx 9/20 = 0.45$$

Null Distribution (empirical)

Give me a number between 1 and 100!



$$\text{P-value} = P(S^n \geq S^0 | \text{1-100 dist})$$

P-values and the null hypothesis

The p-value is the probability of observing your test statistic or a statistic more extreme than it given that it was drawn from the null distribution.

The p-value says NOTHING about the alternate hypothesis

P-value: Important parts

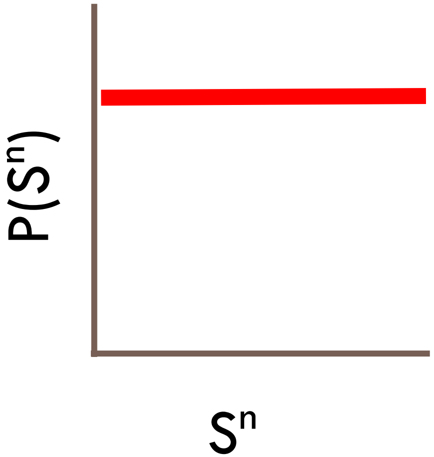
$$\text{P-value} = P(S^n \text{ more extreme than } S^0 | \text{null-dist})$$

<u>S</u>^{O}: The test statics, evaluated for how extreme it is.

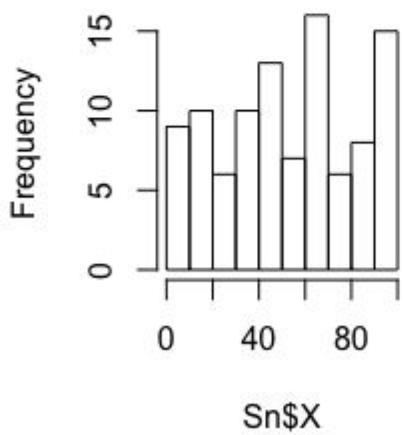
<u>S</u>^{n}: A random variable generated from the null distribution.

<u>Null distribution</u>: The distribution of observed or expected S^n given that the null hypothesis is true.

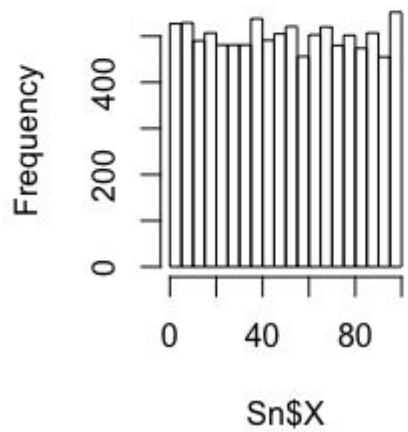
Null distribution: Empirical vs Assumed



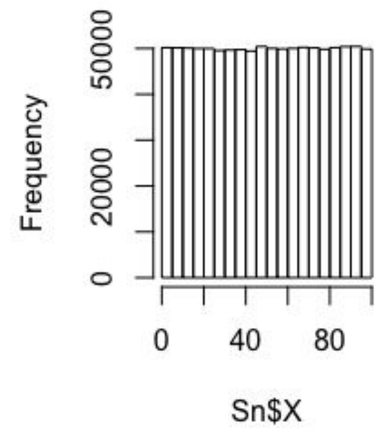
Draws=100



Draws=10,000



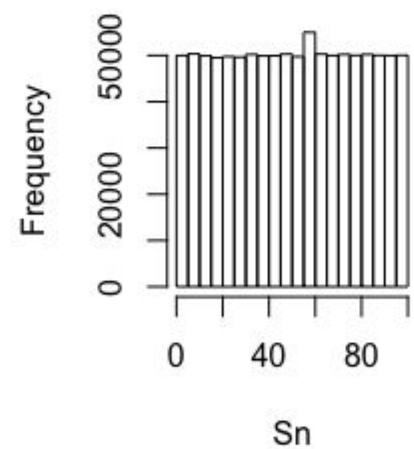
Draws=1,000,000



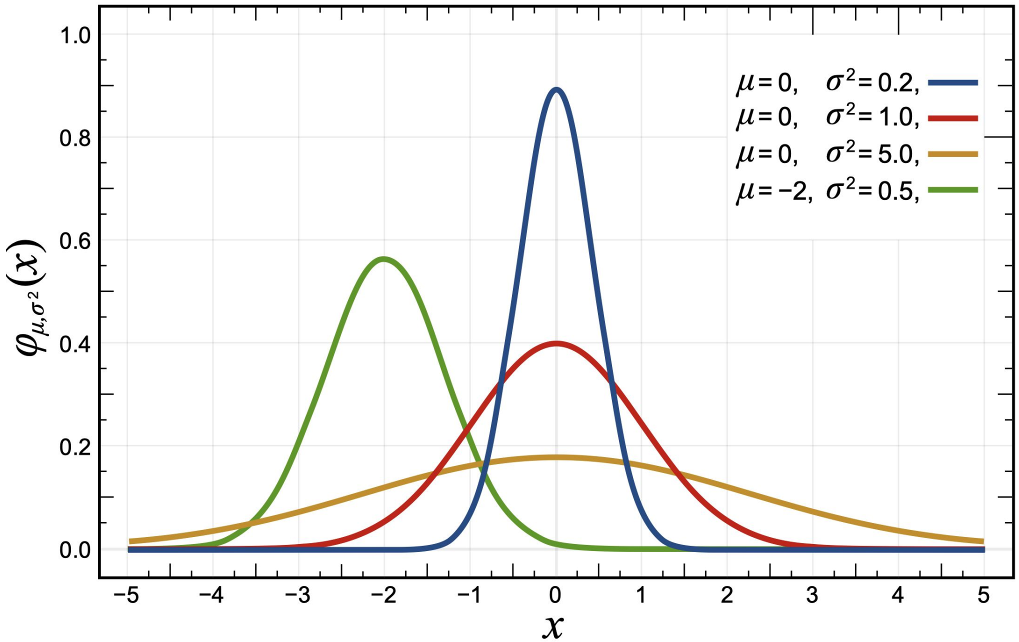
Assumed

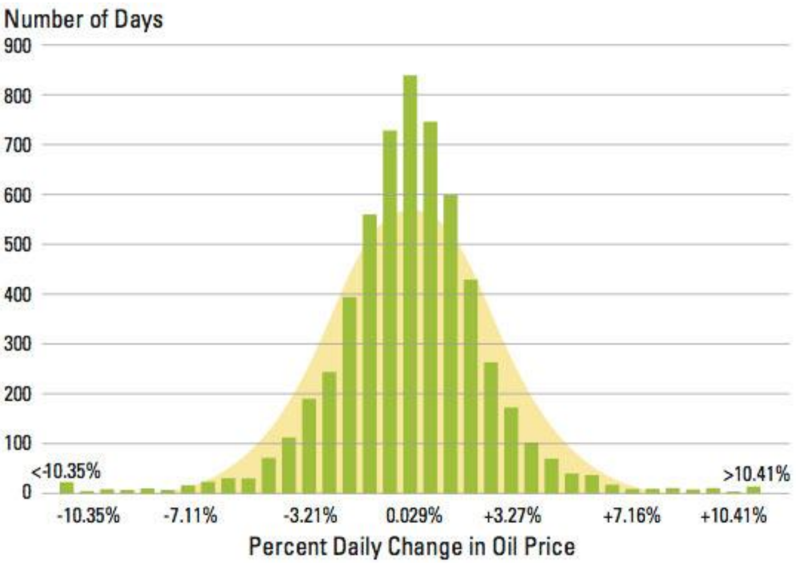
Real

Draws=1,000,000

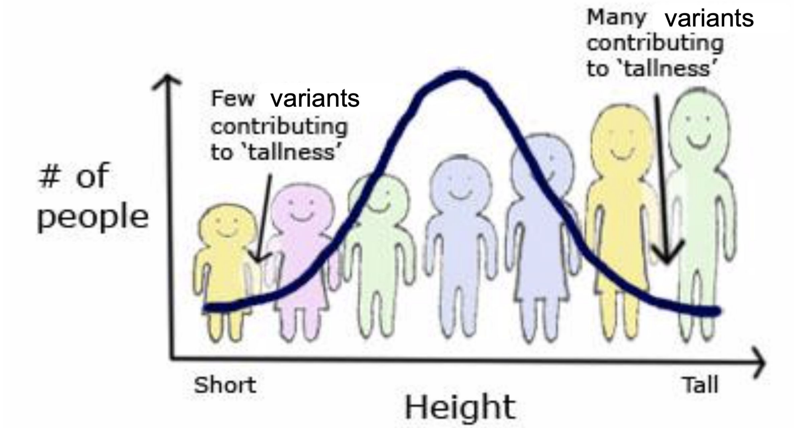


Our favorite null distribution: The normal distribution





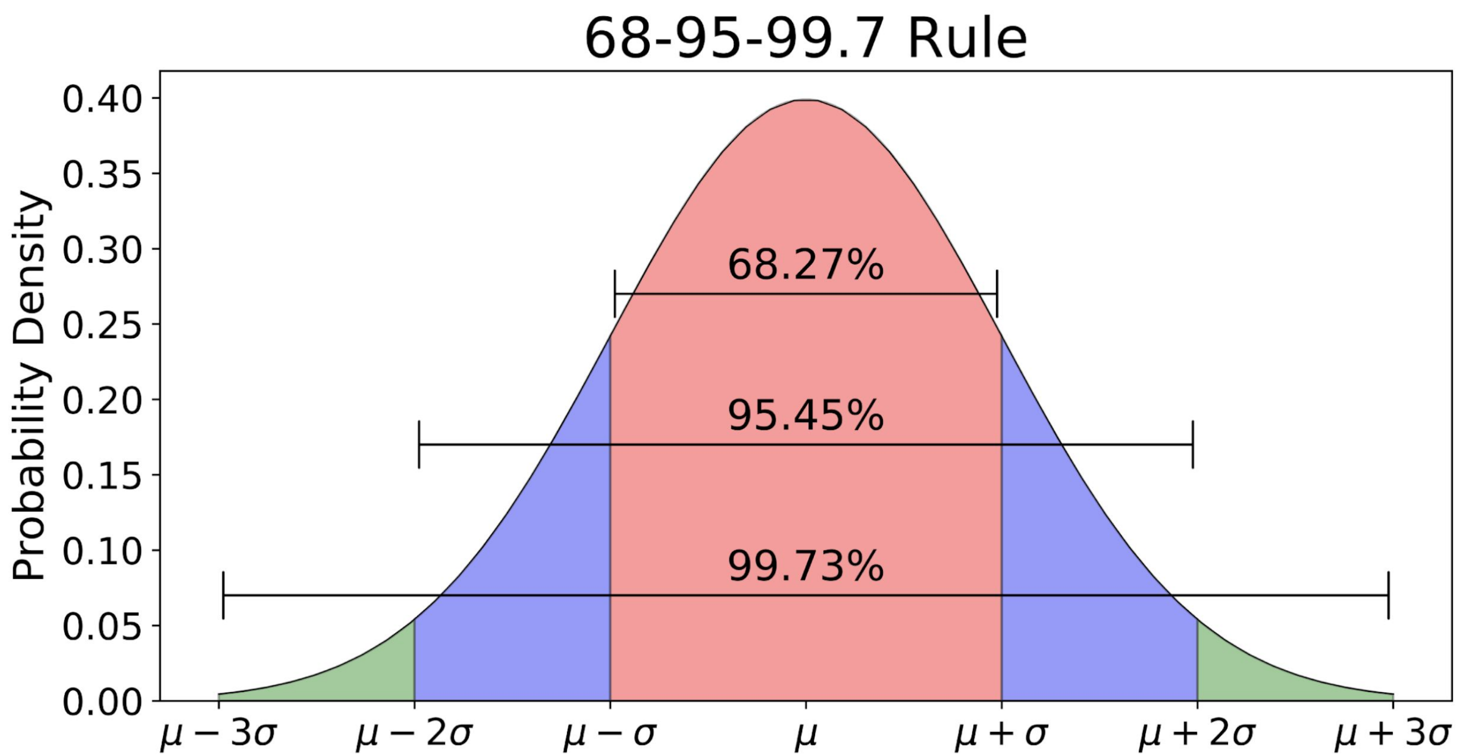
$$Y = X\beta$$



https://marieandmorgan420yo.weebly.com/real-life-examples.html
https://en.wikipedia.org/wiki/Normal_distribution

Our favorite null distribution: The normal distribution

$$Y = X\beta$$



Michael Galarnyk:
https://towardsdatascience.com/understanding-the-68-95-99-7-rule-for-a-normal-distribution-b7b7cbf760c2

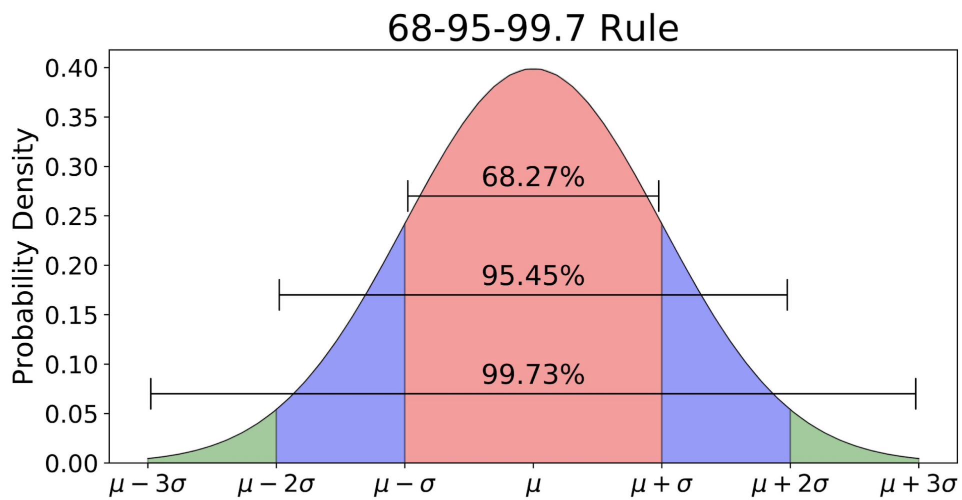
Alternate Hypothesis

$$Y = X_{42}\beta_{42}$$

Height

SNP 42

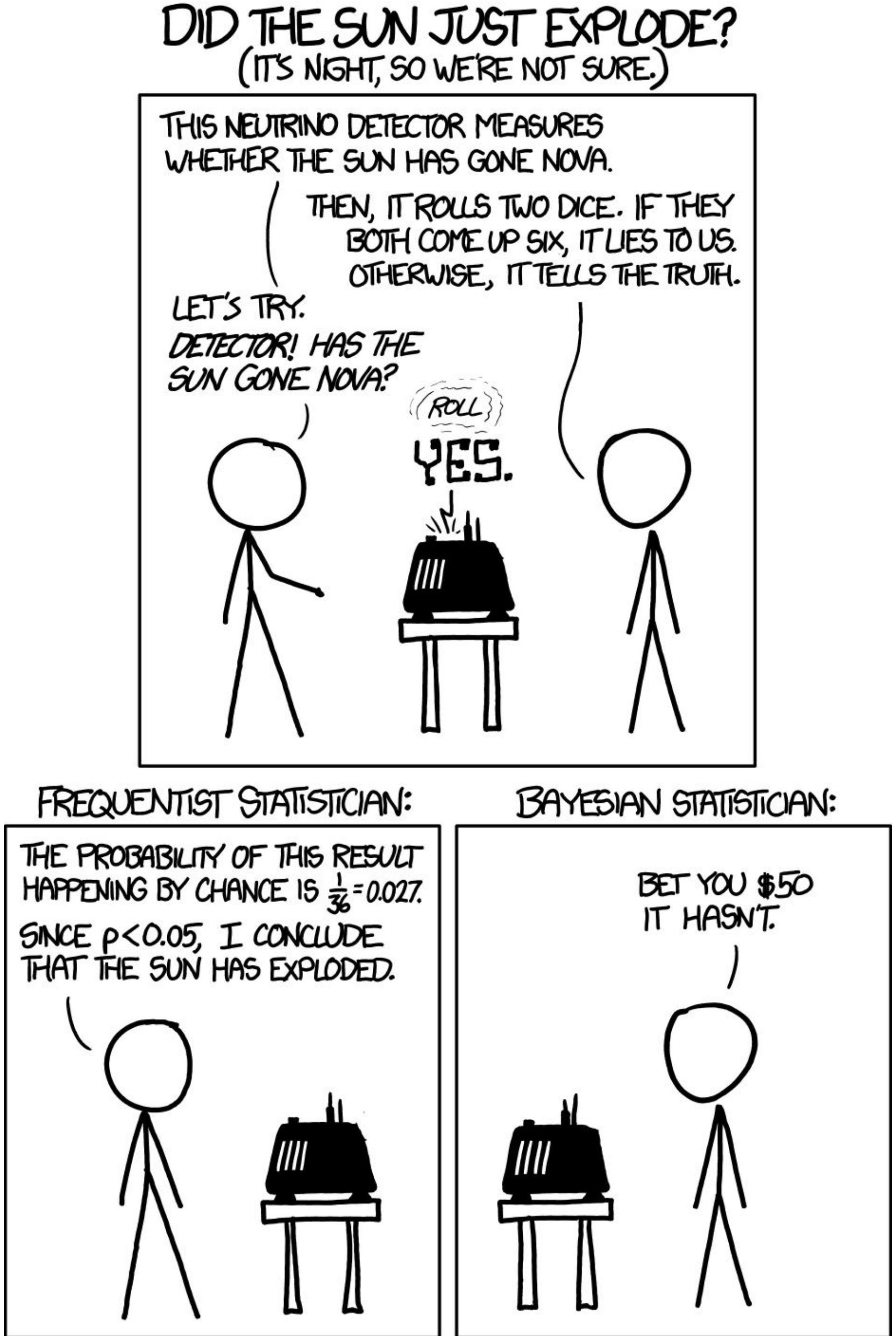
Effect of an alternate allele at SNP 42



Null: $\beta_{42} = 0$

Alt: $\beta_{42} \neq 0$

Michael Galarnyk: https://towardsdatascience.com/understanding-the-68-95-99-7-rule-for-a-normal-distribution-b7b7cbf760c2



No: Sun did not explode and 2 sixes were not rolled

No: Sun exploded!! and two sixes were rolled

Yes: Sun did not explode and 2 sixes were rolled

Yes: Sun exploded!! and two sixes were not rolled

https://xkcd.com/1132/

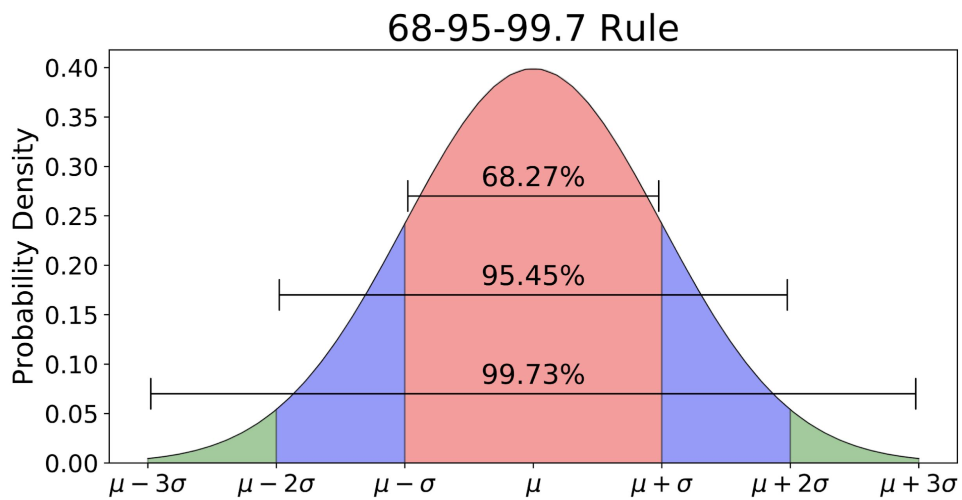
Alternate Hypothesis

$$Y = X_{42}\beta_{42}$$

Height

SNP 42

Effect of an alternate allele at SNP 42



Null: $\beta_{42} = 0$

Alt: $\beta_{42} \neq 0$

Alt: SNP 43 in perfect LD with SNP with non-zero effect

Alt: SNP 43 correlates/tags population structure

Michael Galarnyk:
https://towardsdatascience.com/understanding-the-68-95-99-7-rule-for-a-normal-distribution-b7b7cbf760c2

Significance Threshold….

If the significance threshold for rejecting the Null hypothesis is 0.05, what does that mean?

- You are willing to take a 1/20 chance that you will incorrectly reject the null on EACH test you do.
- If you do 100 test, all of which are truly under the null, you expect 5 "false positives"

True/False Positives/Negatives

True Model

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{18} \\ g_{21} & g_{22} & \cdots & g_{28} \\ \vdots & \vdots & \vdots & \vdots \\ g_{N1} & g_{N2} & \cdots & g_{N8} \end{bmatrix} \begin{bmatrix} 0.2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1.3 \\ 0 \\ -0.05 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix}$$

$$\hat{\beta} = \begin{bmatrix} \boxed{0.19} \\ 0.03 \\ -0.01 \\ -0.07 \\ \boxed{0.12} \\ \boxed{1.6} \\ 0.06 \\ -0.08 \\ 0.01 \\ 0.003 \end{bmatrix} \quad \text{p-values} = \begin{bmatrix} 0.003 \\ 0.34 \\ 0.68 \\ 0.82 \\ 0.04 \\ 0.00003 \\ 0.12 \\ 0.06 \\ 0.21 \\ 0.49 \end{bmatrix}$$

$$Y = X\beta$$

Threshold=0.05	No Effect	Effect
Reject Null	1	2
Accept Null	6	1

Threshold=0.001	No Effect	Effect
Reject Null	0	1
Accept Null	7	2

False Discovery Rate

Threshold=0.05	No Effect	Effect
Reject Null	1	2
Accept Null	6	1

Threshold=0.001	No Effect	Effect
Reject Null	0	1
Accept Null	7	2

$$\mathrm{FDR} = \frac{1}{1+2} = 0.33$$

$$\mathrm{FDR} = \frac{0}{0+1} = 0$$

$$\mathrm{FDR} = \frac{FP}{P} = \frac{FP}{FP+TP}$$

Family-wise Error Rate

Threshold=0.05	No Effect	Effect
Reject Null	1	2
Accept Null	6	1

Threshold=0.001	No Effect	Effect
Reject Null	0	1
Accept Null	7	2

Type 1 Error

$$\mathrm{FWER} = P(\ \mathrm{T1E}\ \geq 1) = 1 - P(\ \mathrm{T1E} = 0)$$

<u>Example Scenario #1</u>
Assume all tests are under null
Threshold is 0.05
You do one test
What is the expected FWER?

$$\mathrm{FWER} = 1 - 0.95 = 0.05$$

$$\mathrm{FWER} = 1 - 0.999 = 0.001$$

<u>Example Scenario #2</u>
Assume all tests are under null
Threshold is 0.05
You do two tests
What is the expected FWER?

$$\mathrm{FWER} = 1 - 0.95^2 = 0.0975$$

$$\mathrm{FWER} = 1 - 0.999^2 = 0.002$$

<u>Example Scenario #3</u>
Assume all tests are under null
Threshold is 0.05
You do M tests
What is the expected FWER?

$$\mathrm{FWER} = 1 - 0.95^M$$

$$\mathrm{FWER} = 1 - 0.999^M$$

Family-wise Error Rate

$$\mathrm{FWER} = P(\ \mathrm{T1E}\ \geq 1) = 1 - P(\ \mathrm{T1E} = 0)$$

<u>Example Scenario #1</u>
Assume all tests are under null
Threshold is 0.05
You do one test
What is the expected FWER?

$$\mathrm{FWER} = 1 - 0.95 = 0.05$$

$$\mathrm{FWER} = 1 - 0.999 = 0.001$$

<u>Example Scenario #2</u>
Assume all tests are under null
Threshold is 0.05
You do two tests
What is the expected FWER?

$$\mathrm{FWER} = 1 - 0.95^2 = 0.0975$$

$$\mathrm{FWER} = 1 - 0.999^2 = 0.002$$

<u>Example Scenario #3</u>
Assume all tests are under null
Threshold is 0.05
You do M tests
What is the expected FWER?

$$\mathrm{FWER} = 1 - 0.95^M$$

$$\mathrm{FWER} = 1 - 0.999^M$$

$$t \approx \frac{FWER}{M}$$

Bonferroni corrected threshold of 0.05

$$t \approx \frac{0.05}{1{,}000{,}000} = 0.00000005 = 5 \times 10^{-8}$$

Conclusions

The p-value is the probability of observing your test statistic or a statistic more extreme than it given that it was drawn from the null distribution.

The p-value says NOTHING about the alternate hypothesis

"Significance Threshold"

- Your likelihood of wrongly rejecting the null hypothesis (type 1 error) on each test assuming the null is true
- Can be used to control the FWER and FDR
- Bonferroni corrected threshold is used to control the FWER