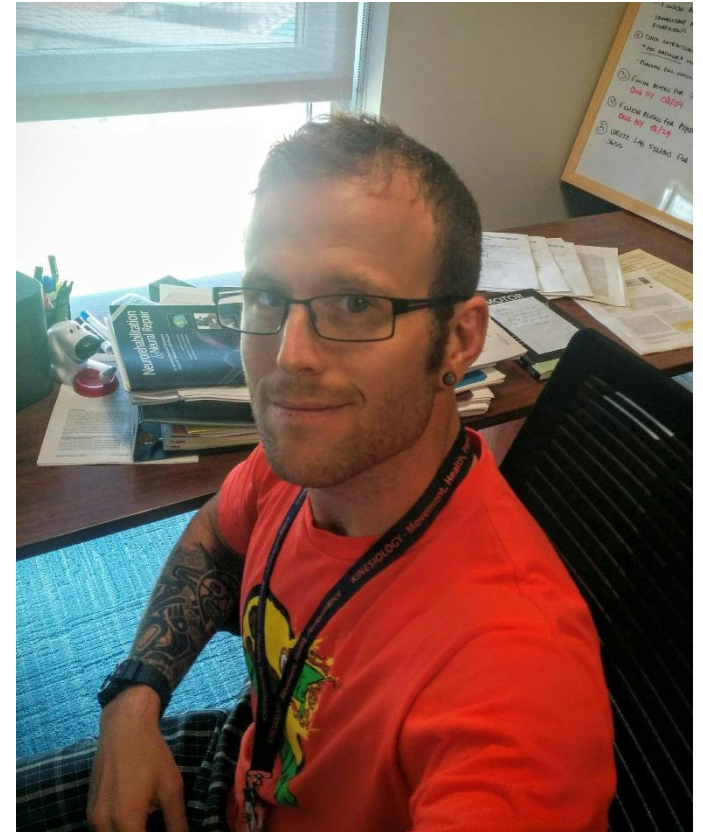


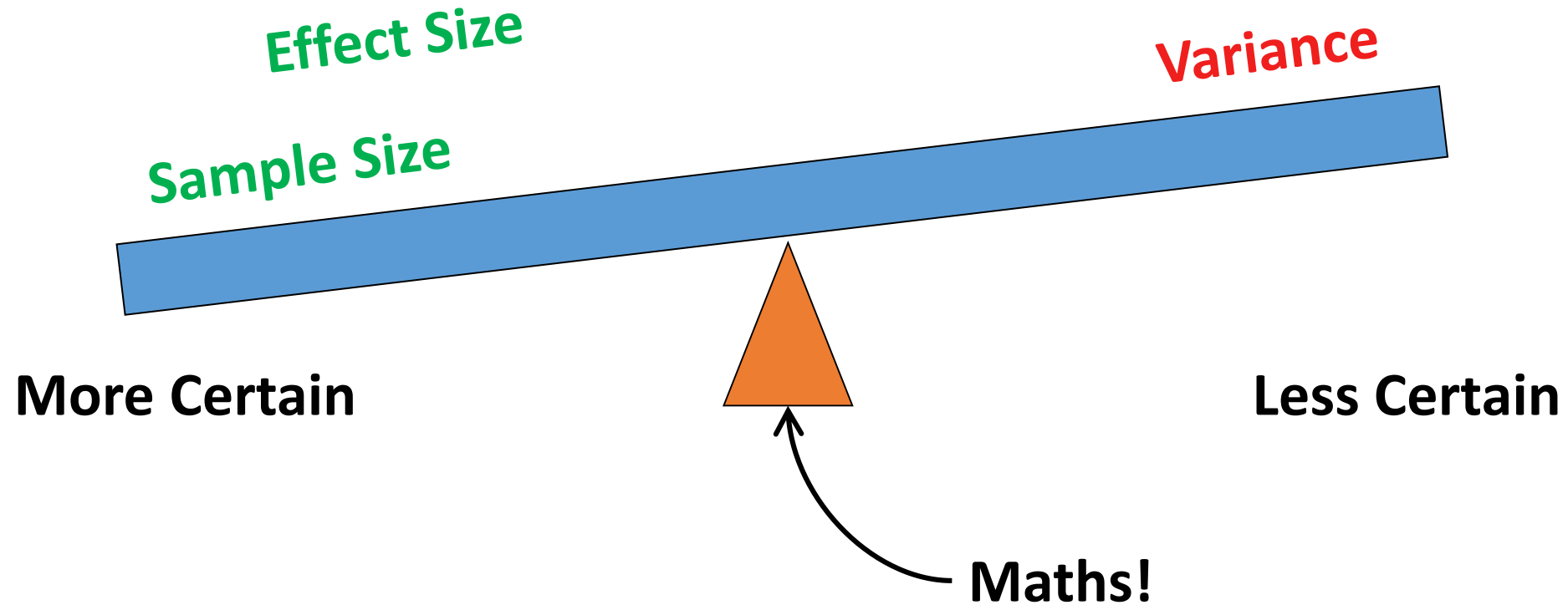
How much power do I need?

Keith Lohse, PhD

[http://compare2what.blogspot.com/
motorlearnlab@gmail.com](http://compare2what.blogspot.com/motorlearnlab@gmail.com)

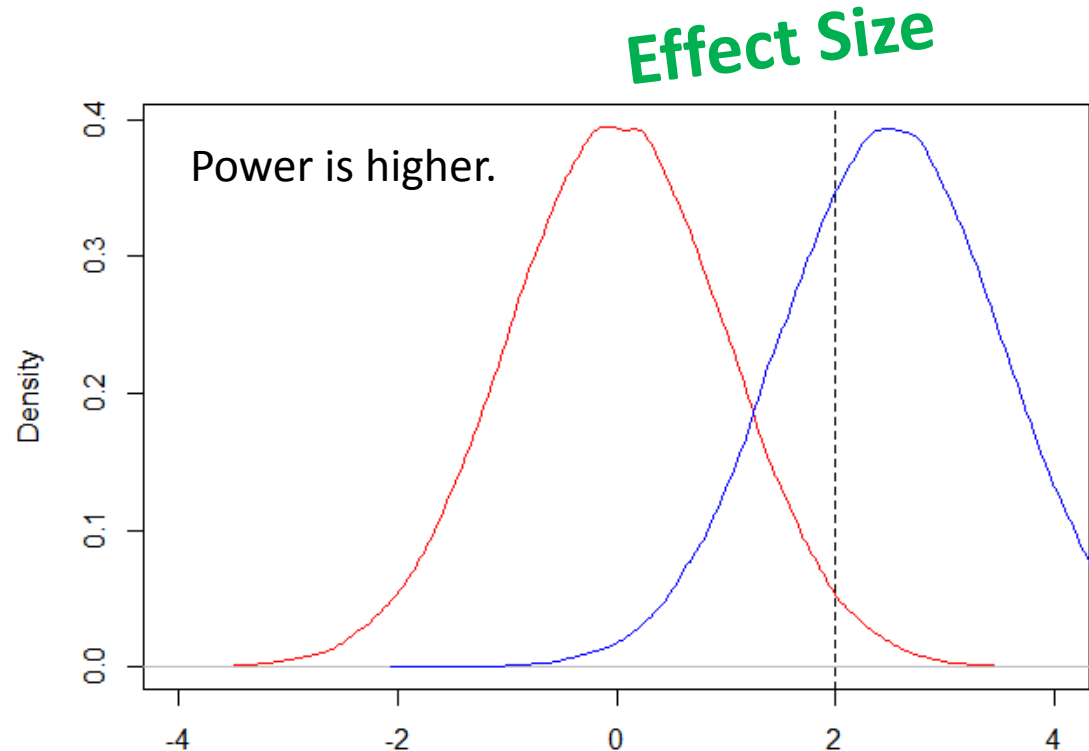
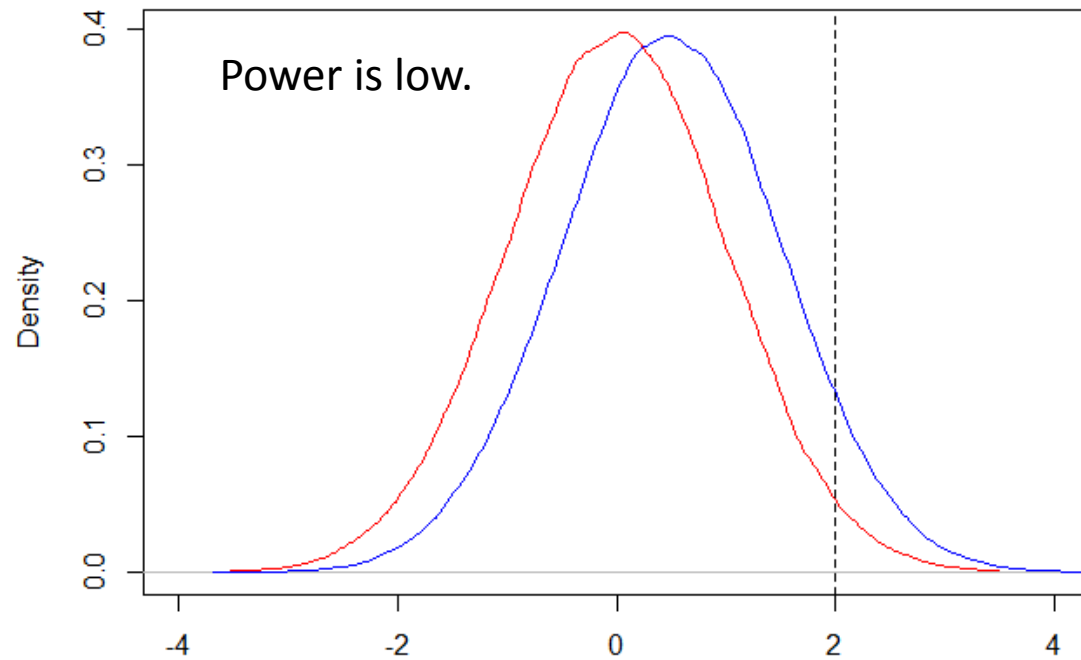


What is statistical power?



What is statistical power?

Distributions of Sample Means

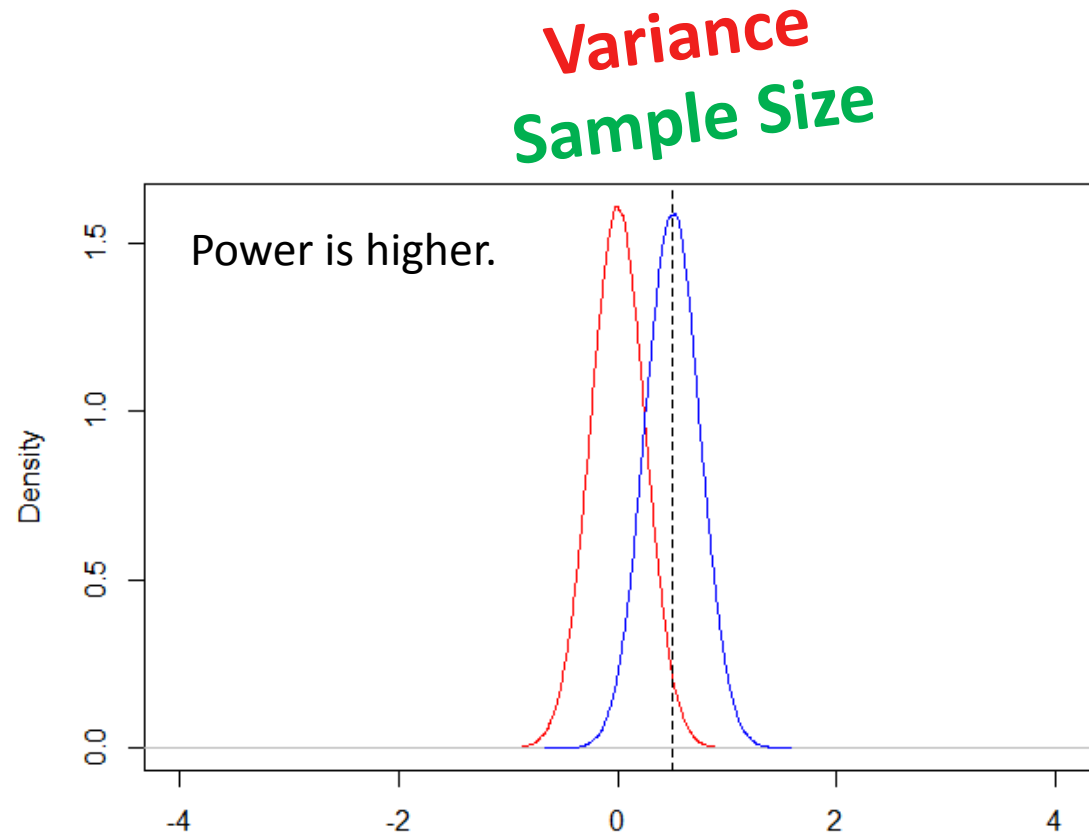
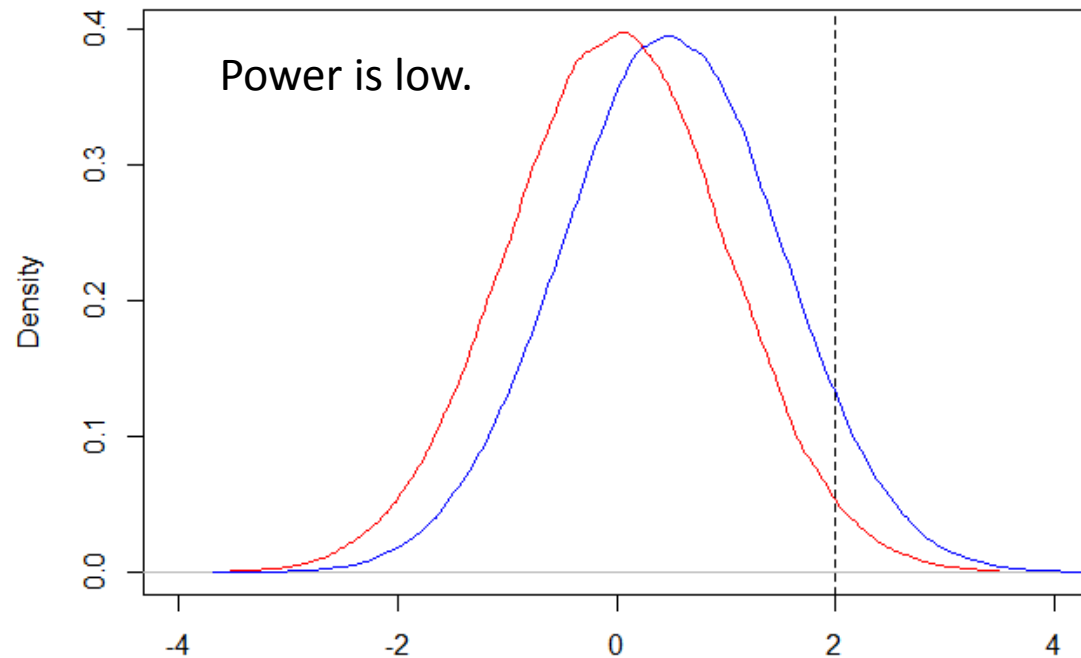


Assuming the Null Hypothesis is true.

Distribution based on the “real” effect.

What is statistical power?

Distributions of Sample Means



Assuming the Null Hypothesis is true.

Distribution based on the “real” effect.

$$SE = \frac{\sigma}{\sqrt{N}} \approx \frac{s}{\sqrt{N}}$$

The “standard error” is the special name for the standard deviation of a distribution of sample means.

How big do our samples need to be to reliably detect an effect?

- Effect size (either ES or SES).
 - Mean difference, correlation, odds-ratio, etc.
- Type 1 error rate, false alarm rate, usually $\alpha = 0.05$.
 - Generally corresponds to $2*SE$ (for normal distributions).
- Type 2 error rate, miss rate, usually $\beta = 0.20$
 - Power = $1-\beta$
 - “How much of the blue curve is past the $2*SE$ critical value?”

Are adult males heavier than adult females?

- How big is the effect?
 - Go to the literature, get yourself an estimate.
 - Adult, white/non-Hispanic males (n = 4406)
 - Mean = 89.4 kg; SD = 26.5 kg
 - Adult, white/non-Hispanic females (n = 4235)
 - Mean = 74.7 kg; SD = 26.0 kg

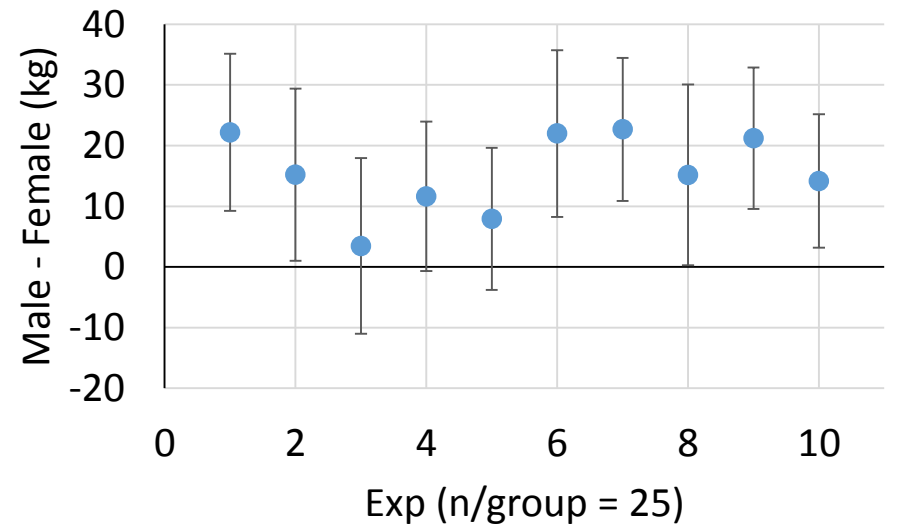
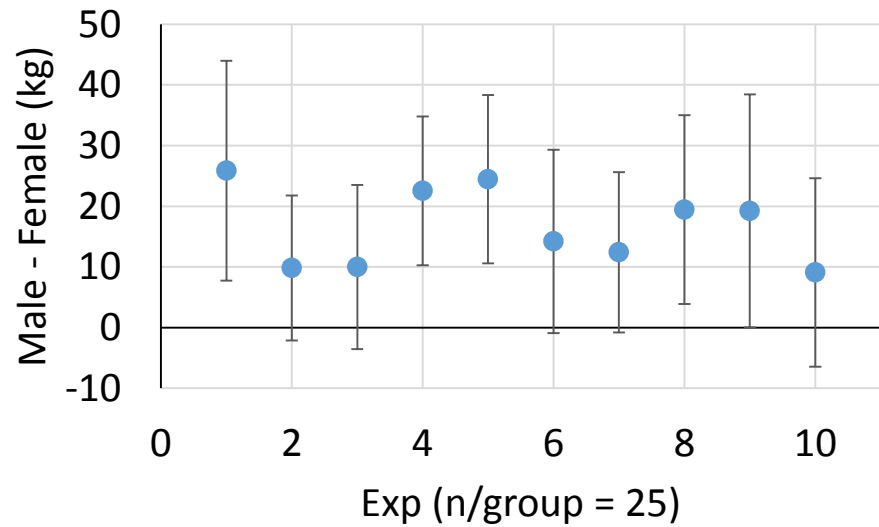
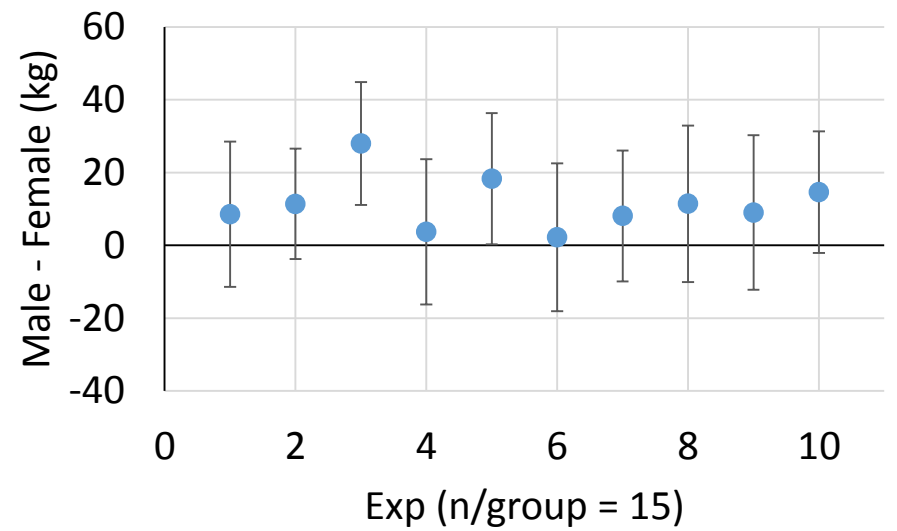
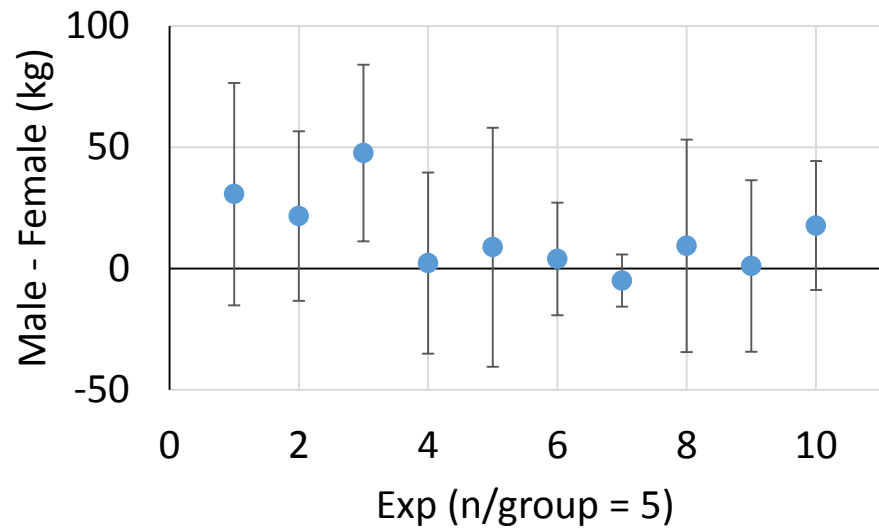
$$ES = 89.4 - 74.7 = 14.7$$

$$SES = d = \frac{m_1 - m_2}{s_p} = \frac{14.7}{26.26} = 0.56$$

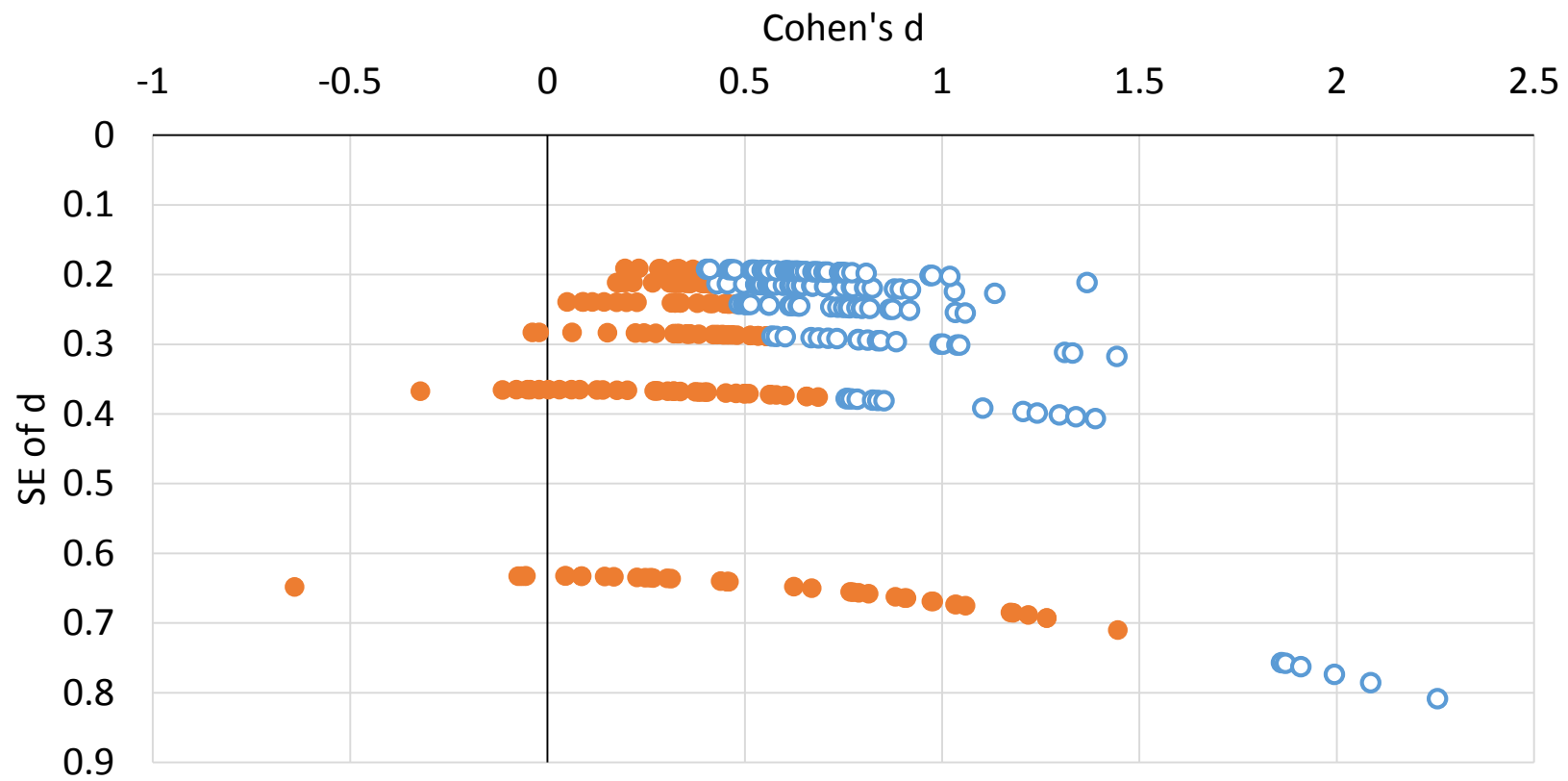
How many subjects per group do we need to reliably show this effect?

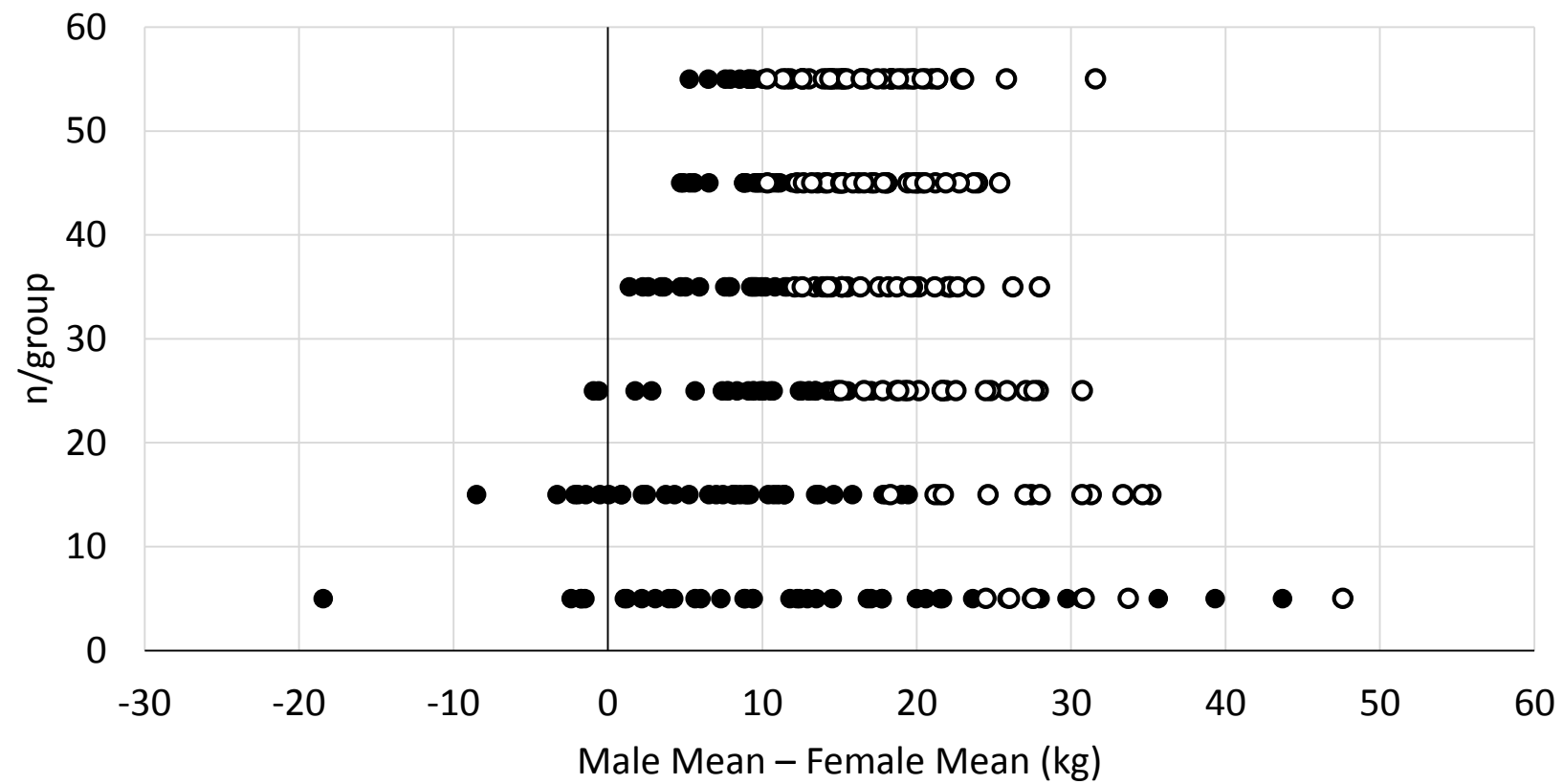
Are adult males heavier than adult females?

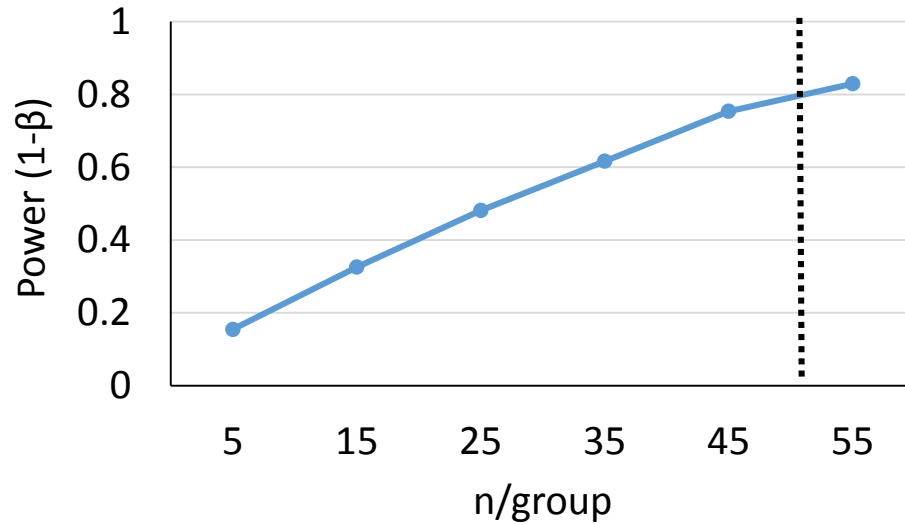
- Set up populations with the parameters 89.4(26.5) and 74.7(26.0).
- Run $k=10,000$ “experiments” with variable n/group {5-55}.
 - Random sampling.
 - Independent samples t-test.
- Plot the results!
 - How big does n/group need to be before we have >80% statistically significant results?



Showing the first 10 of $k = 10,000$ random simulations.







In order to reliably* detect a difference in weight between adult males and females in the US, you need approximately 50 people per group.

*Reliable operationally defined as 80% power.

Based on the proportion of significant results $k = 10,000$ simulations.

- Are your effects bigger than sex differences in weight?
 - Regardless, you should probably consider running fewer, larger experiments.
- What else can I do to increase my power?
 - Intelligent design: repeated measures designs, covariates, multiple baselines.
 - Maximize ES.

Thanks for watching!

- Thanks also to Jim McDonald at Auburn University.
- The weight example was chosen from a talk entitled “Life After P -Hacking” by Uri Simonsohn, Leif Nelson, and Joe Simmons, check it out!
- If you have any questions, you can email them to me at motorlearnlab@gmail.com or leave a comment below.

