Power Calculations for Categorical Factors

Ryan Lekivetz
Senior Research Statistician Developer
JMP

Joint work with Luyao Lin & Derek Bingham
Simon Fraser University
Burnaby, BC, Canada
Power Calculations for the Real World

• We’ve seen some of the “messy” cases, but what about the simpler case of linear regression?
  • minimum effect size to be detected
  • significance level of the hypothesis test
  • variability
  • design characteristics

• For two-level factors, the effect size is directly related to the coefficients in a least squares regression model - easy to calculate power

• We’ll use $\Delta$, max-min over design range
  • $\Delta=2$ corresponds to coefficient of 1
Continuous Factor

Difference between max and min from low to high settings of X

\[ \Delta = 2 \]
2-level Categorical Factor

Difference between max and min from the two different levels of $X$

$\Delta = 2$
Power for Categorical Factors

For number of levels: $k > 2$

Usually interested in the overall effect rather than individual parameters.

$H_0$: All treatments (levels) have the same effect
$H_a$: Not all treatment effects are the same

“Effect Tests” in Fit Model or
“Effect Power” in Custom Design

What does $\Delta$ mean with $k > 2$ levels?
Notes

• Error/RMSE/sigma = 1
  • This way we only need to talk about $\Delta$ in terms of max-min.
• Fix $\alpha=0.05$
• Categorical effects will use the “sum to zero” constraint.
  • For $k$ levels, only specify $k-1$ coefficients.
$k = 3$

Which is the best choice for $\Delta = 2$?
\( k = 4, \Delta = 2 \)
\[ k = 6, \Delta = 2 \]
\[ k = 3, \ N = 24, \ \Delta = 2 \]

Power calculations:

- \text{power} = 0.924
- \text{power} = 0.977
$k = 4, \ N=24, \ \Delta = 2$

power = .755

power = .802

power = .912

power = .973
$k = 6, \ N=24, \ \Delta = 2$

power = .432

power = .586

power = .674

power = .923
Why does it matter?

Consequences of using the worst-case

• Use additional runs to achieve desired power.
• Drop levels out of consideration from a categorical factor.
• May not run the experiment at all.
What does $\Delta = 2$ mean for $k>2$?

- Does the worst-case seem plausible?
- There are infinitely many choices once we fix the max and min.
- What if we think about simulating possible configurations for the effects?
Treatment effects

• Let \( \boldsymbol{\tau}^a = (\tau_1^a, \tau_2^a, \ldots, \tau_k^a)' \) be vector of treatment effects under \( H_a \).

• Set

\[
\begin{align*}
\tau_k^a - \tau_1^a &= \Delta \\
\tau_1^a &\leq \tau_2^a \leq \ldots \leq \tau_k^a
\end{align*}
\]

• Assign multivariate distribution \( f_\tau(\cdot) \) with domain \([\tau_1^a, \tau_k^a]\).

• Reflects the belief of the effects in \( \boldsymbol{\tau}^a \).
Sum-to-zero constraint

• Easy to modify the vector of treatment effects.
• Let $\tau^{a*}$ be the constrained vector, generated from

$$(\tau_1^a - \delta, \tau_2^a - \delta, \ldots, \tau_i^a - \delta, \ldots, \tau_k^a - \delta)'$$

where

$$\delta = \sum_i \frac{\tau_i^a}{k}$$

• Then we can generate the vector and adjust.
Power Function

\[ \eta(D, \tau^a) = P(\text{reject } H_0 | \hat{H}_a : \tau = \tau^{a*} \text{ is true}) \]

- Random variable for design \( D \), under domain of \( f_{\tau}(\cdot) \).
- Can simulate from \( f_{\tau}(\cdot) \) to assess the power.
Benefits of the Power Function

Allows us to ask questions

• What is the average power?

\[ E[\eta(D, \tau^a)] = \int \eta(D, \tau^a)f_{\tau}(\tau^a)d\tau^a. \]

• What is the probability that the power is at least 80%?

• What is the .05 quantile of the power distribution?
  • 95% of the time the power is greater than what number?

Off to JMP...
Power for categorical factors is not so simple.

Only looking to the worst-case can be costly and not a practically likely situation.

We can use the idea of simulating the treatment effects to get a distribution of the power.

Think about power for how you do the analysis.
Thank you!

jmp.com