Multiple Comparisons and False Positives

Set a family size and alpha level below to see how quickly the chance of at least one false positive grows, and how much a Bonferroni correction brings it back down.

StatReason

Multiple Comparisons and False Positives

Number of comparisons (family size)

5 comparisons

Per-comparison alpha

With 5 comparisons, uncorrected false positive risk is 22.6 percent; Bonferroni correction brings it to 4.9 percent.

Chance of at least one false positive, uncorrected	22.6%
Bonferroni-corrected alpha per comparison	0.0100
Chance of at least one false positive, corrected	4.9%

Apply this to your own comparisons: **Bonferroni Correction Calculator**

What this answers

This tool answers "if I run many statistical tests, how likely am I to get at least one false positive purely by chance, and how does a correction fix that?" Every individual test run at alpha .05 has a 5 percent false positive rate on its own, but running many tests multiplies the opportunities for at least one of them to be a false positive.

Why the family-wise error rate grows

If each of several independent comparisons has a 5 percent chance of a false positive, the chance that none of them is a false positive is 0.95 raised to the power of the number of comparisons. The chance of at least one false positive is 1 minus that value, which grows quickly: with just 10 independent comparisons at alpha .05, the chance of at least one false positive is already above 40 percent.

How a Bonferroni correction controls it

A Bonferroni correction divides the target overall alpha by the number of comparisons and uses that smaller value as the per-comparison threshold, which keeps the overall family-wise false positive rate close to the original target regardless of how many comparisons you run. This control comes at a cost: each individual comparison becomes harder to call significant, which can reduce power to detect genuine differences, especially with a large number of comparisons.

Worked example

Set family size to 10 and alpha to .05 above. The uncorrected chance of at least one false positive across all 10 comparisons is well above 40 percent. The Bonferroni-corrected per-comparison alpha becomes .005, and the corrected overall chance of at least one false positive returns close to the original 5 percent target.

Assumption audit

> **Calculated from your data:** in this visual, the exact false-positive probabilities from your chosen family size and alpha; the linked calculator applies the same correction to your own set of p-values.

> **Evidence to review:** whether your comparisons are truly independent, since the exact formula used here assumes independence and correlated comparisons behave somewhat differently.

> **You must verify:** whether your set of comparisons was planned before seeing the data or selected afterward, since a correction applied only to comparisons chosen after seeing the results does not fully address the underlying problem.

Source

This explanation follows the multiple-comparisons guidance in the NIST/SEMATECH e-Handbook of Statistical Methods and the shared statistical reasoning contract every StatReason engine is built against.

Limitations

This visual assumes independent comparisons for the uncorrected calculation; Tukey's HSD and Dunn's test use methods specifically designed for correlated comparisons and can be less conservative than Bonferroni in those cases.