Simulation as Estimation

Power by Simulation

$$
\begin{aligned}
P_n(T > c) &= E[I(T > c)] \\
&= \sum_t I(t > c) P_n(T = t) \\
&= \sum_{\mathbf{x}} I(t(\mathbf{x}) > c) \mathcal{P}_n(\mathbf{X} = \mathbf{x}) \\
&\approx \frac{1}{M} \sum_{i=1}^{M} I(t(\mathbf{X}_i) > c)
\end{aligned}
$$

Advantages

- For normal linear models, not much use except to check your work.
- For simple multinomial models, Agresti's formulas for the non-centrality parameter work well for large samples.
- For less familiar models, simulation can be faster and easier than trying to do it analytically.
- Can be less error prone, too.

Monte Carlo Integration (or Summation)

$$\int g(x)\,dx \quad = \quad \int \frac{g(x)}{f(x)}\,f(x)\,dx$$

$$\frac{1}{M}\sum_{i=1}^{M}\frac{g(X_i)}{f(X_i)} \quad \stackrel{a.s.}{\longrightarrow} \quad \int g(x)\,dx$$

Statistical considerations apply

- Large-sample theory (definitely)
- Central limit theorem
- Confidence intervals
- Tests
- Factorial designs
- Variance reduction: Make g(x)/f(x) as close to a constant as possible to reduce the variance of the estimator.

A Toy Example

$$\int_0^1 x^4 \left| \sin(100 \tan(\cos(\frac{1}{1-x}))) \right| dx$$