Power calculations in R

Outline:

- Power calculations for z tests (pair of means)
  - Central and non-central χ^2 distributions
  - Confidence intervals, hypothesis testing and power calculations for a variance
  - Non-central F distributions
- Power calculations for fixed-effects ANOVA (differences in a group of means)
- Power calculations for random-effects ANOVA (differences in variances)

Normal Distribution calculations in R

Let U = unit normal. Prob(U < x), use pnorm(x)

Examples:

Prob(U < 1.0). pnorm(1.0) returns 0.8413

Prob(U > 2.3). 1- pnorm(2.3) returns 0.01072

Critical values. To find x where Prob(U < x) = p, qnorm(p)

Examples:

Find x where Prob(U < x) = 0.999

qnorm(0.999) returns 3.0902

Find x such that Prob(|U| < 0.99)

qnorm(0.995) returns 2.576

Check: Prob(|U| < 2.576) = Prob(U < 2.576) - Prob(< -2.576)

pnorm(2.576)-pnorm(-2.576) returns 0.990005

Normal tests and Power

Define $z_{(\square)}$ = Prob(U < $z_{(\square)}$) = $\square$. Note this is qnorm($\square$)

Likewise, note that Pr(U > $z_{(\square)}$) = 1 - $\square$ = Pr (U < $z_{(1-\square)}$).

Example: $\square$ = 0.95. 1 - $\square$ = 0.05

Prob(U > $z_{(0.05)}$) = (U < $z_{(0.95)}$) = qnorm(0.95)

R returns 1.644

Suppose we have n observations from a normal with mean μ_0 and variance σ^2.

Hence, $\bar{x}$ is a normal with mean μ_0 and variance σ^2/n, so that $(x - \mu_0)/(\sigma^2/n)^{1/2} \sim U$ (a unit normal)

For a one-sided test with (say) $\alpha = 0.01$, want is the critical value $T_C(\alpha)$? Namely, for what value is $\Pr(x > T_C(\alpha)) = \alpha$ (=0.01 in our example).

$$
\begin{aligned}
\Pr(x > T_C(\alpha)) &= \Pr(x - \mu_0 > T_C(\alpha) - \mu_0) \\
&= \Pr([x - \mu_0]/s > [T_C(\alpha) - \mu_0)]/s) \\
&= \Pr(U > [T_C(\alpha) - \mu_0)]/s) = \alpha
\end{aligned}
$$

Thus $[T_C(\alpha) - \mu_0)]/s = z_{(1-\alpha)}$ $\qquad s = (\sigma^2/n)^{1/2}$

$$T_C(\alpha) = \mu_0 + s\, z_{(1-\alpha)} = \mu_0 + (\sigma^2/n)^{1/2} z_{(1-\alpha)}$$

$$T_C(\alpha) = \mu_0 + s\, z_{(1-\alpha)} = \mu_0 + (\sigma^2/n)^{1/2}\, z_{(1-\alpha)}$$

In our case, $\alpha = 0.01$, $z_{(1-\alpha)} = z_{(0.99)}$ = qnorm(0.99) = 2.33

So, if our null hypothesis is that the mean = 10 and $\sigma^2 = 4$, our critical 1% (one-sided) test for sample size n is just $10 + (4/n)^{1/2} * 2.33$

Defining functions in R

We want to compute the critical value for different values of n, so let's define a function in R to do this

```
Nc <- function(n) 10+sqrt(4/n)*2.33
```

More generally, let's let a vary as well

```
Ncrit <- function(n,a) 10+sqrt(4/n)*qnorm(1-a)
```

$T_C(\alpha) = \mu_0 + s\, z_{(1-\alpha)} = \mu_0 + (\sigma^2/n)^{1/2}\, z_{(1-\alpha)}$

```
Ncrit <- function(n,a) 10+sqrt(4/n)*qnorm(1-a)
```

Ncrit(15,0.01) returns 11.20

Ncrit(200,0.001) returns 10.43

Most generally, we can write a function also allow μ_0 and σ^2 to vary,

```
Ncrit <- function(m0,var,n,a) m0+sqrt(var/n)*qnorm(1-a)
```

Ncrit(10,4,15,0.01) returns 11.20 (as expected)

Ncrit(10,8,15,0.01) returns 11.70

Power for Means tests

Now suppose the true distribution of x is that it has mean μ_1 (with the same variance σ^2). The power of this test is the probability that the same mean exceeds the critical value, $T_c(\alpha) = \mu_0 + (\sigma^2/n)^{1/2} z_{(1-\alpha)}$

Here, the mean is now normal with μ_1, σ^2, and the power is

$$\begin{aligned} Pr(x > T_C(\alpha)) &= Pr(x - \mu_1 > T_C(\alpha) - \mu_1) \\ &= Pr([x - \mu_1]/s > [T_C(\alpha) - \mu_1)]/s) \\ &= Pr(U > [T_C(\alpha) - \mu_1)]/ (\sigma^2/n)^{1/2}) \end{aligned}$$

We can program this in R

```
power <- function(m0,m1,var,a,n)   {
temp <-(Ncrit(m0,var,n,a)- m1)/sqrt(var/n);
1-pnorm(temp)}
```

Mean: null mean = 10, true mean = 11, var = 4, α = 0.01

power(10,11,4,0.01,20) returns 0.464

Suppose we increase sample size to 60

power(10,11,4,0.01,200) returns 0.939

curve(power(10,11,4,0.01,x),10,200)

Central and Non-central χ^2 distributions

Central χ^2 distributions

Recall that if $x_i \sim N(0,1)$, then $(x_1 + x_2 \cdots + x_n) \sim \chi^2_n$

If $x_i \sim N(0,\sigma^2)$,

$$\sum_{i=1}^{n} x_i^2 \sim \sigma^2 \chi^2_n \qquad \sum_{i=1}^{n} (x_i - \overline{x})^2 \sim \sigma^2 \chi^2_{(n-1)}$$

Noncentral χ^2 distributions

If $x_i \sim N(\mu_i,\sigma^2)$,

$$\sum_{i=1}^{n} x_i^2 \sim \sigma^2 \chi^2_{n,\lambda} \quad \text{with} \quad \lambda = \sum_{i=1}^{n} \mu_i^2$$

Chi-square with n degrees of freedom and noncentrality parameter λ.

Likewise, for $x_i \sim N(\mu_i, \sigma^2)$,

$$\sum_{i=1}^{n} (x_i - \overline{x})^2 \sim \sigma^2 \chi^2_{(n-1),\lambda} \quad \text{with} \quad \lambda = \sum_{i=1}^{n} \frac{\mu_i^2}{\sigma^2}$$

Central & noncentral χ^2 in R

$\Pr(\chi^2_n < x)$ = pchisq(x,n)

Example: Compute $\Pr(\chi^2_{35} < 41.5)$

pchisq(41.5,35) returns 0.792

$\Pr(\chi^2_{n,l} < x)$ = pchisq(x,n,l)

Example: Compute $\Pr(\chi^2_{35,6} < 41.5)$

pchisq(41.5,35,6) returns 0.551

To find x such that $Pr(\chi^2_n < x) = a$, use qchisq(a,n)

Example: Compute the 95% value for a χ^2_{10}.

qchisq(0.95,10) returns 18.31

To find x such that $Pr(\chi^2_{n,l} < x) = a$, use qchisq(a,n,l)

Example: Compute the 95% value for a $\chi^2_{10,3.5}$.

qchisq(0.95,10,3.5) returns 24.27

Confidence limits on variance estimators

If $z_i \sim N(\mu, \sigma^2)$, i.e., they all have a common mean, we recover a central χ^2 distribution,

$$S = \sum_{i=1}^{n} (x_i - \overline{x})^2 \sim \sigma^2 \chi^2_{n-1}$$

Since S is very closely related to the sample variance, we can obtain confidence intervals, critical values, and compute power for simple variance estimates.

$$Var = \frac{S}{n-1} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2 \sim \frac{\sigma^2}{n-1} \chi^2_{n-1}$$

We designate appropriate χ^2 values, let $\Pr(\chi^2_n \leq \chi^2_{n,[\alpha]}) = \alpha$

First, note that

$$\Pr(\chi^2_{n-1,[\alpha/2]} \leq \chi^2_{n-1} \leq \chi^2_{n-1,[1-\alpha/2]}) = 1 - \alpha$$

Recalling that the scaled sample variance follows a χ^2

$$\Pr\left(\chi^2_{n-1,[\alpha/2]} \leq \frac{(n-1)Var}{\sigma^2} \leq \chi^2_{n-1,[1-\alpha/2]}\right) = 1-\alpha$$

Further noting that for a < x < b, that 1/a > 1/x > 1/b gives

$$\Pr\left(\frac{1}{\chi^2_{n-1,[\alpha/2]}} \geq \frac{\sigma^2}{(n-1)Var} \geq \frac{1}{\chi^2_{n-1,[1-\alpha/2]}}\right) = 1-\alpha$$

Thus our 1-a level confidence interval for the true variance given the sample variance is

$$\Pr\left(\frac{(n-1)}{\chi^2_{n-1,[\alpha/2]}} Var \geq \sigma^2 \geq \frac{(n-1)}{\chi^2_{n-1,[1-\alpha/2]}} Var\right) = 1-\alpha$$

$$\Pr\left(\frac{(n-1)}{\chi^2_{n-1,[\alpha/2]}} Var \geq \sigma^2 \geq \frac{(n-1)}{\chi^2_{n-1,[1-\alpha/2]}} Var \right) = 1-\alpha$$

Example: suppose n = 20 and our sample variance is 10. What is a 99% confidence for the true variance in this case?

First note that here □ = 1.

- qchisq(0.005,19) gives $\chi^2_{10,[0.005]} = 6.85$
- qchisq(0.995,19) gives $\chi^2_{10,[0.995]} = 38.58$

Lower limit = 19*10/38.58 = 4.9

Upper limit = 19*10/6.85 = 27.76

R code.

Note that c(x,y) returns an array, 1st element = lower 2nd element = upper

```
chiCI<-function(var,n,a) {
 low<-qchisq(a/2,n-1);
  upper<-qchisq(1-a/2,n-1);
  c((n-1)*n*var/upper,(n-1)*var/low)}
```

chiCI(10,50,0.001) returns

5.55 21.50

Critical Values and Power for One-sided tests

Consider a null of $\sigma^2 = \sigma^2_0$ vs the alternative $\sigma^2 > \sigma^2_0$

What is the critical value $C(n, \alpha)$ under the alternative to give an error (Type I error) of α?

$$\Pr(Var > C(n,\alpha)) = \alpha$$ This is what we want

$$\Pr\left(Var\frac{(n-1)}{\sigma_0^2} > C(n,\alpha)\frac{(n-1)}{\sigma_0^2}\right) = \alpha$$

$$\Pr\left(\chi^2_{n-1} > C(n,\alpha)\frac{(n-1)}{\sigma_0^2}\right) = \alpha$$

$$\chi^2_{n-1,[1-\alpha]} = C(n,\alpha)\frac{(n-1)}{\sigma_0^2}$$

Hence, the critical value becomes

$$C(n,\alpha) = \chi^2_{n-1,[1-\alpha]} \frac{\sigma_0^2}{n-1}$$

Suppose the true variance is □$^2{}_1$. What is the probability that we reject the null (the power)? This is the prob that Var exceeds *C*:

$$\Pr(Var > C(n,\alpha) \,|\, \sigma_1^2) = \Pr\left(Var \frac{(n-1)}{\sigma_1^2} > C(n,\alpha) \frac{(n-1)}{\sigma_1^2} \right)$$

Which reduces to a very simple form:

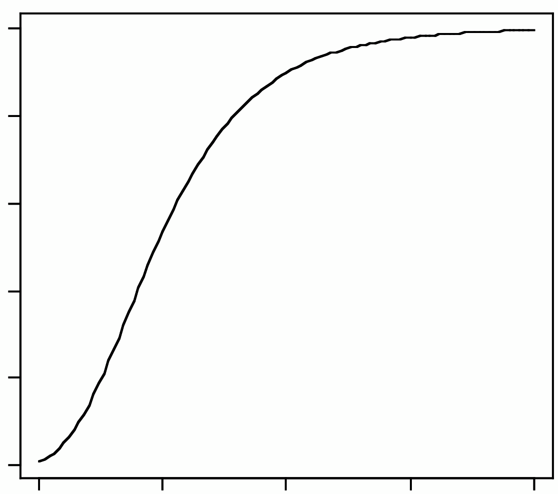
$$\Pr\left(\chi^2_{n-1} > \frac{\sigma_0^2}{\sigma_1^2} \chi^2_{n-1,[1-\alpha]} \right)$$

R Code

```
chipower<-function(var,n,alpha) 1-pchisq(qchisq(1-alpha,n)/var,n)
```

σ^2_0 / σ^2_1 Sample size

Example: plot of power as a function of the variance ratio for n = 20, alpha = 0.01. `curve(chipower(x,20,0.01),1,20)`



The critical value and power for the other one-sided test,
a null of $\sigma^2 = \sigma^2_0$ vs the alternative $\sigma^2 < \sigma^2_0$

Critical value

$$C(n,\alpha) = \chi^2_{n-1,[\alpha]} \frac{\sigma_0^2}{n-1}$$

Reject null when Var < $C(n,\alpha)$

Power:

$$\Pr\left(\chi^2_{n-1} < \frac{\sigma_0^2}{\sigma_1^2} \chi^2_{n-1,[\alpha]} \right)$$

Central and Non-central F distributions

Central F distributions

$$(\chi^2_i/i)\,/\,(\chi^2_k/k) \sim F_{i,k}$$

Ratio of two chi-squares
divided by their df's

F with numerator i
and denominator
k degrees of
freedom

Non-central F distributions

$$(\chi^2_{i,\lambda}/i)\,/\,(\chi^2_k/k) \sim F_{i,k,\lambda}$$

Ratio of non-central over
Central chi-squares

F with numerator i,
denominator
k degrees of freedom
and non-centrality
parameter λ.

F Distribution values in R

$\Pr(F_{i,k} < x)$ pf(x,i,k) $\Pr(F_{i,k,lam} < x)$ pf(x,i,k,lam)

Example: Compute $\Pr(F_{20,30} > 1.5)$

1-pf(1.5,20,30) returns 0.153

Example: Compute $\Pr(F_{20,30,2} > 1.5)$

1-pf(1.5,20,30,2) returns 0.215

Find x such that $\Pr(F_{i,k} < x) = a$ qf(a,i,k)

Example: Compute the 95% critical value for an F with 12 numerator and 25 denominator dfs

qf(0.95,12,25) returns 2.16

Power of F-tests

The idea is the same as with normal tests

- Assign a critical value *C* based on the null
- Compute the probability that the test statistic exceed *C* given the assumed alternative values

- Fixed-effects ANOVA:

  Null: treatment effects $\square = 0$, so $y_i \sim N(0, \square^2_e)$.
  Alternative: some $\square$ not zero, so $y_i \sim N(\square, \square^2_e)$.
  Non-zero means generate non-central F

- Random-effects ANOVA:

  Null: treatment variance $\square^2_{\square} = 0$, so $y_i \sim N(0, \square^2_e)$.
  Alternative: $\square^2_{\square} > 0$, so $y_i \sim N(0, \square^2_e + \square^2_{\square})$.
  Non-zero variance generates scaled F (i.e., $\square$*F)

Fixed-effects: One-way ANOVA

$$y_{ij} = u + \tau_i + e_{ij}$$

N fixed factors, n replicates/factor. The test statistic is

$$f = \frac{\mathrm{MS}_t}{\mathrm{MS}_e} = \frac{\mathrm{SS}_t/(N-1)}{\mathrm{SS}_e/N(n-1)}$$

The distribution of the numerator follows since

$\overline{y}_i \sim \mathrm{N}(\tau_i, \sigma_e^2/n)$ which gives

$$\mathrm{SS}_t \sim n(\sigma_e^2/n)\chi^2_{N-1,\lambda} \quad \textbf{with} \quad \lambda = \sum_{i=1}^{N} \frac{{\tau_i}^2}{\sigma_e^2/n}$$

We can think of this as a variance-like term,

$$\sigma_\tau^2 = \frac{1}{N-1}\sum_{i=1}^{N} {\tau_i}^2 \qquad \lambda = n\,(N-1)\,\frac{\sigma_\tau^2}{\sigma_e^2}$$

The error sums of squares is distributed at

$$\mathrm{SS}_e \sim \sigma_e^2 \chi^2_{N(n-1)}$$

Giving the distribution of the test statistic f as

$$f \sim \left(\frac{\sigma_e^2}{\sigma_e^2}\right) \left[\frac{\chi^2_{N-1,\lambda}/(N-1)}{\chi^2_{N(n-1)}/N(n-1)} \right] \sim F_{N-1,N(n-1),\lambda}$$

Under the null hypothesis of no treatment effect, □ = 0 and f is centrally F distributed. Hence the a-level critical value for the null is

$$f > F_{(N-1),N(n-1),[1-\alpha]}$$

Under the alternative, $\square = n(N-1)\square^2_{\square}/\square^2_e$, and the power of the test is simply $\Pr\left(F_{N-1,N(n-1),\lambda} > F_{N-1,N(n-1),[1-\alpha]}\right)$

$\square = n(N-1)\square^2_{\square}/\square^2_e$, power is

$$\Pr\left(F_{N-1,N(n-1),\lambda} > F_{N-1,N(n-1),[1-\alpha]}\right)$$

Example: Suppose the treatment "variance" is 10% of the total variance and we have N = 5 groups. For a test of $\square = 0.001$, what n is needed for 80% power?

The critical value becomes $F_{4,5*n-1,[0.999]}$ = qf(0.999,4,5*(n-1))

What is the noncentrality parameter?

$\square^2_{\square}/(\square^2_e + \square^2_{\square}) = 1/(\square^2_e/\square^2_{\square} + 1) = 0.1$, implying $\square^2_e/\square^2_{\square} = 9$, so that
$\square = n(N-1)\square^2_{\square}/\square^2_e = n(4/9) = 0.444*n$

```
power <-function(n) {
crit <- qf(0.999,4,5*(n-1));
1-pf(crit,4,5*(n-1),0.444*n)}
```

Let's plot power as a function of n

```
curve(power(x),20,100)
```

$\square = n(N-1)\square^2_{\square}/\square^2_e$, power is

$$\Pr\left(F_{N-1,N(n-1),\lambda} > F_{N-1,N(n-1),[1-\alpha]}\right)$$

General expression

```
power <-function(N,n,var,alpha) {
crit <- qf(1-alpha, N-1, N *(n -1));
ncp <- n * (N-1)*var;
1-pf(crit, N-1,N *(n -1),ncp)}
```

$\square^2_{\square}/\square^2_e$ versus $\square^2_{\square}/(\square^2_e + \square^2_{\square})$

A few brief comments on the ratio of treatment variance to error variance ($\square^2_{\square}/\square^2_e$) vs. the fraction of total variance accounted for by the treatment effects ($\square^2_{\square}/(\square^2_e + \square^2_{\square})$), which is the r^2 for the simple one-way ANOVA (so for simplicity, we will refer to this ratio as r^2.

Let $x = \square^2_{\square}/\square^2_e$, then we can express r^2 in terms of x, and vise-versa, Specifically,

$$r^2 = \square^2_{\square}/(\square^2_e + \square^2_{\square}) = 1/(\square^2_e /\square^2_{\square} + 1) = 1/(1/x + 1) = x/(1+x)$$

Likewise, $x = r^2/(1-r^2)$, giving $\square = n(N-1)\square^2_{\square}/\square^2_e = n(N-1)[r^2/(1-r^2)]$

Example: Suppose $\square^2_{\square}/\square^2_e = 0.2$. Then $r^2 = 0.2/(1+0.2) = 0.167$

Example: if $r^2 = 0.05$, what is $\square^2_{\square}/\square^2_e$ and $\square$?

$$\square^2_{\square}/\square^2_e = 0.05/(1-0.05) = 0.053, \quad \square = n(N-1)\square 0.053$$

Power under random effects design

$$y_{ij} = u + \square_i + e_{ij}$$

f test as before, distribution of error sum of squares SS_e as before. However, since $\square \sim N(0, \square^2_{\square})$, for the treatment sum of squares, $\mathrm{SS}_t \sim (n\,\sigma_t^2 + \sigma_e^2)\chi^2_{N-1}$

The resulting distribution of the test statistic becomes

$$f = \frac{\mathrm{SS}_t/(N-1)}{\mathrm{SS}_e/N(n-1)} \sim \frac{(n\,\sigma_t^2 + \sigma_e^2)\chi^2_{N-1}/(N-1)}{\sigma_e^2\chi^2_{N(n-1)}/N(n-1)}$$

$$\sim \left(1 + n\,\frac{\sigma_t^2}{\sigma_e^2}\right) F_{(N-1),N(n-1)}$$

Hence,

$$\frac{F}{1 + n\,(\sigma_t^2/\sigma_e^2)} \sim F_{N-1,N(n-1)}$$

When the treatment variance is zero (null hypothesis), same critical value as for fixed-effects model.

When the treatment variance is non-zero, the power is

$$\Pr\left[F_{N-1,N(n-1)} > \frac{F_{N-1,N(n-1),[1-\alpha]}}{1 + n\,(\sigma_t^2/\sigma_e^2)} \right]$$

Note that, as with the fixed-effects case, we can replace the variance ratio by $r^2/(1-r^2)$

$$\Pr\left[F_{N-1,N(n-1)} > \frac{F_{N-1,N(n-1),[1-\alpha]}}{1 + n\left(\sigma_t^2/\sigma_e^2\right)} \right]$$

Now the (R) code for power is

```
rpower <-function(N,n,var,alpha) {
crit <- qf(1-alpha, N-1, N *(n -1));
temp <- 1+n*var;
1-pf(crit/temp, N-1,N *(n -1))}
```

What about differences in power for fixed vs. random?

```
curve(power(5,x,0.111,0.001)-rpower (5,x,0.111,0.001),5,100)
```