In statistical hypothesis testing, the α-risk (probability of rejecting true null) is controlled by construction, but the β-risk (probability of accepting false null) is left open. To also control the β-risk *at specific alternatives*, one needs large enough sample sizes.

The probabilty of rejecting the null is known as the power of the test; it is the α-risk when the null is true, and it is one minus the β-risk when the null is false.

The power is a property of specific tests *as procedures*, constructed for specific null and alternative hypotheses in specific settings. *Power analysis is logically detached from the execution of the tests using observed data*, though past data might assist power analysis for future studies in similar settings.

Power analysis is "analytically" possible in limited settings, of which some are covered in these notes. Power analysis is always doable via simulations, but with *one sample size, one alternative at a time*.

1 One/Two-Sample Settings

Power analysis is analytically tractable for one/two-sample tests concerning normal means, but numerical calculations require variance values which are typically unknown. Ballpark variance estimates might be obtained from past/pilot studies or using empirical ranges taken as 4σ or 6σ; larger σ values are conservative for sample size planning.

1.1 One-Sample Tests

Observing $Y_i \sim N(\mu, \sigma^2)$, $i = 1, \dots, n$, one has $\bar{Y} \sim N(\mu, \frac{\sigma^2}{n})$ independent of $(n-1)s^2/\sigma^2 \sim \chi^2_{n-1}$, leading to $\frac{\bar{Y}-\mu}{s/\sqrt{n}} \sim t_{n-1}$.

Consider the one-sided hypotheses $H_0 : \mu \le \mu_0$ vs. $H_a : \mu > \mu_0$. For σ^2 known, one uses a z-test, rejecting H_0 when $\frac{\bar{Y}-\mu_0}{\sigma/\sqrt{n}} > z_{1-\alpha}$. For σ^2 unknown, one uses a t-test, rejecting H_0 when $\frac{\bar{Y}-\mu_0}{s/\sqrt{n}} > t_{1-\alpha,n-1}$.

At $\mu > \mu_0$, $\frac{\bar{Y}-\mu_0}{\sigma/\sqrt{n}}$ has mean $\frac{\mu-\mu_0}{\sigma/\sqrt{n}}$ and unit variance, and $\frac{\bar{Y}-\mu_0}{s/\sqrt{n}}$ follows a non-central t-distribution with df $(n-1)$ and noncentrality parameter $\frac{\mu-\mu_0}{\sigma/\sqrt{n}}$. The power of the z-test is given by

$$p(\mu) = P\big(\tfrac{\bar{Y}-\mu_0}{\sigma/\sqrt{n}} > z_{1-\alpha}\big) = P\big(\tfrac{\bar{Y}-\mu}{\sigma/\sqrt{n}} > -\tfrac{\mu-\mu_0}{\sigma/\sqrt{n}} + z_{1-\alpha}\big) = 1 - \Phi\big(-\tfrac{\mu-\mu_0}{\sigma/\sqrt{n}} + z_{1-\alpha}\big). \quad (1)$$

The power of the t-test can be obtained in `R` using function `pt`.

As a toy example, consider $\frac{\mu-\mu_0}{\sigma} = 0.8$ and $n = 15$. At level $\alpha = 0.05$, the following `R` code calculates the powers of the z-test and the t-test.

```
1-pnorm(qnorm(.95)-0.8*sqrt(15))
1-pt(qt(.95,14),14,0.8*sqrt(15))
```

To ensure $p(\mu) \ge 1-\beta$ for $\mu \ge \mu_1$ with the z-test, one needs $z_{1-\alpha} - \sqrt{n}\frac{\mu_1-\mu_0}{\sigma} \le z_\beta = -z_{1-\beta}$, or

$$n \ge \big(\tfrac{z_{1-\alpha}+z_{1-\beta}}{\delta}\big)^2, \quad (2)$$

where $\delta = \frac{\mu_1 - \mu_0}{\sigma}$ is the effect size. The sample size for t-test should be slightly bigger, obtainable via trial-and-error using the `pt` function. With $\delta = 0.8$, $\alpha = 0.05$, and $\beta = 0.1$, try

```
((qnorm(.95)+qnorm(.9))/0.8)^2
1-pt(qt(.95,13),13,0.8*sqrt(14))
1-pt(qt(.95,14),14,0.8*sqrt(15))
```

For the two-sided hypotheses $H_0 : \mu = \mu_0$ vs. $H_a : \mu \neq \mu_0$, the z-test rejects H_0 when $\left|\frac{\bar{Y}-\mu_0}{\sigma/\sqrt{n}}\right| > z_{1-\alpha/2}$, and the t-test rejects H_0 when $\left|\frac{\bar{Y}-\mu_0}{s/\sqrt{n}}\right| > t_{1-\alpha/2,n-1}$. The power of the z-test at $\mu \neq \mu_0$ is given by

$$p(\mu) = 1 - P\big(-z_{1-\alpha/2} \leq \tfrac{\bar{Y}-\mu_0}{\sigma/\sqrt{n}} \leq z_{1-\alpha/2}\big) = 1 - \Phi\big(-\tfrac{\mu-\mu_0}{\sigma/\sqrt{n}} + z_{1-\alpha/2}\big) + \Phi\big(-\tfrac{\mu-\mu_0}{\sigma/\sqrt{n}} - z_{1-\alpha/2}\big). \quad (3)$$

With $|\mu - \mu_0| = 0.8\sigma$, $n = 15$, and $\alpha = 0.05$, the powers of the z-test and t-test are via

```
1-pnorm(qnorm(.975)-0.8*sqrt(15))+pnorm(-qnorm(.975)-0.8*sqrt(15))
1-pt(qt(.975,14),14,0.8*sqrt(15))+pt(-qt(.975,14),14,0.8*sqrt(15))
```

By symmetry, $p(\mu_0 + \gamma) = p(\mu_0 - \gamma)$, and for $\mu > \mu_0$ and n reasonably large, the last term in (3) is negligible, so to ensure $p(\mu) \geq 1 - \beta$ for $|\mu - \mu_0| \geq |\mu_1 - \mu_0|$ with the z-test, one needs

$$n \geq \big(\tfrac{z_{1-\alpha/2} + z_{1-\beta}}{\delta}\big)^2, \quad (4)$$

where $\delta = \frac{|\mu_1 - \mu_0|}{\sigma}$. For $\delta = 0.8$, $\alpha = 0.05$, and $\beta = 0.1$, try

```
((qnorm(.975)+qnorm(.9))/0.8)^2
1-pnorm(qnorm(.975)-0.8*sqrt(17))+pnorm(-qnorm(.975)-0.8*sqrt(17))
1-pt(qt(.975,17),17,0.8*sqrt(18))+pt(-qt(.975,17),17,0.8*sqrt(18))
1-pt(qt(.975,18),18,0.8*sqrt(19))+pt(-qt(.975,18),18,0.8*sqrt(19))
```

1.2 Two-Sample Tests

With independent samples $Y_{ij} \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2$, $j = 1, \ldots, n_i$, $\bar{Y}_1 - \bar{Y}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$, and z-tests can always be used for σ_1^2, σ_2^2 known.

For $\sigma_1^2 = \sigma_2^2$ possibly unknown, $(n_1 + n_2 - 2)s_p^2 = (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 \sim \sigma^2\chi^2_{n_1+n_2-2}$, and $\frac{\bar{Y}_1 - \bar{Y}_2}{s_p\sqrt{1/n_1 + 1/n_2}}$ follows a t-distribution with df $(n_1 + n_2 - 2)$ and noncentrality $\frac{\mu_1 - \mu_2}{\sigma\sqrt{1/n_1 + 1/n_2}}$.

For $\sigma_1^2 \neq \sigma_2^2$ unknown, the Behrens-Fisher problem is analytically intractable.

Consider the two-sided hypotheses $H_0 : \mu_1 = \mu_2$ vs. $H_a : \mu_1 \neq \mu_2$. Assuming $\sigma_1^2 = \sigma_2^2$, $n_1 = n_2$, the z-test rejects H_0 when $\left|\frac{\bar{Y}_1 - \bar{Y}_2}{\sigma\sqrt{2/n}}\right| > z_{1-\alpha/2}$, and the t-test rejects when $\left|\frac{\bar{Y}_1 - \bar{Y}_2}{s_p\sqrt{2/n}}\right| > t_{1-\alpha/2,2(n-1)}$. The power of the z-test is seen to be

$$p(\mu_1 - \mu_2) = 1 - \Phi(-\tfrac{\mu_1 - \mu_2}{\sigma\sqrt{2/n}} + z_{1-\alpha/2}) + \Phi(-\tfrac{\mu_1 - \mu_2}{\sigma\sqrt{2/n}} - z_{1-\alpha/2}). \quad (5)$$

With $|\mu_1 - \mu_2| = 0.8\sigma$, $n = 15$, and $\alpha = 0.05$, the powers of the z-test and t-test are available via

```
1-pnorm(qnorm(.975)-0.8*sqrt(15/2))+pnorm(-qnorm(.975)-0.8*sqrt(15/2))
1-pt(qt(.975,28),28,0.8*sqrt(15/2))+pt(-qt(.975,28),28,0.8*sqrt(15/2))
```

To ensure $p(\mu_1-\mu_2)\geq 1-\beta$ for $|\mu_1-\mu_2|\geq d$ with the z-test, one needs

$$n\geq 2\big(\tfrac{z_{1-\alpha/2}+z_{1-\beta}}{d/\sigma}\big)^2. \tag{6}$$

With $d=0.8\sigma$, $\alpha=0.05$, and $\beta=0.1$, try

```
2*((qnorm(.975)+qnorm(.9))/0.8)^2
1-pnorm(qnorm(.975)-0.8*sqrt(33/2))+pnorm(-qnorm(.975)-0.8*sqrt(33/2))
1-pt(qt(.975,64),64,0.8*sqrt(33/2))+pt(-qt(.975,64),64,0.8*sqrt(33/2))
1-pt(qt(.975,66),66,0.8*sqrt(34/2))+pt(-qt(.975,66),66,0.8*sqrt(34/2))
```

1.3 Paired Tests

Observing $Y_{ij}=\mu_i+\beta_j+\epsilon_{ij}$, $i=1,2$, $j=1,\dots,n$, $\epsilon_{ij}\sim N(0,\sigma^2)$, one is to test $H_0:\mu_1=\mu_2$ vs. $H_a:\mu_1\neq\mu_2$. Working with $d_j=Y_{1j}-Y_{2j}=\mu_1-\mu_2+e_j$, where $e_j=\epsilon_{1j}-\epsilon_{2j}\sim N(0,2\sigma^2)$, the z-test rejects H_0 when $\big|\frac{\bar{d}}{\sigma\sqrt{2/n}}\big|>z_{1-\alpha/2}$, and the t-test rejects H_0 when $\big|\frac{\bar{d}}{s_d/\sqrt{n}}\big|>t_{1-\alpha/2,n-1}$, where $s_d^2=\frac{1}{n-1}\sum_{j=1}^n(d_j-\bar{d})^2$.

The power and sample size for the z-test are as given in (5) and (6), but the meaning of σ here differs from that in §1.2. Also, the t-test here has df $(n-1)$ instead of $2(n-1)$.

Note that β_j's do not appear in the paired tests so can be arbitrary. Assuming $b_j\sim N(\mu_b,\sigma_b^2)$, independent of ϵ_{ij}, the marginal variance of Y_{ij} is $\tilde{\sigma}^2=\sigma_b^2+\sigma^2$ and the correlation between Y_{1j} and Y_{2j} is $\rho=\frac{\sigma_b^2}{\sigma_b^2+\sigma^2}=\frac{\tilde{\sigma}^2-\sigma^2}{\tilde{\sigma}^2}$, yielding $\sigma^2=\tilde{\sigma}^2(1-\rho)$.

2 One-Way ANOVA

Consider $Y_{ij}\sim N(\mu_i,\sigma^2)$, $i=1,\dots,a$, $j=1,\dots,n_i$. One has $\bar{Y}_i\sim N(\mu_i,\sigma^2/n_i)$ and $(N-a)s_p^2=\sum_{i=1}^k(n_i-1)s_i^2\sim\chi^2_{N-a}$, where $N=\sum_{i=1}^a n_i$. When $a=2$, this reduces to §1.2 with $\sigma_1^2=\sigma_2^2$.

Multiple pairs of (H_0,H_a) could be formed to address different aspects of interest in applications, and power analysis for the respective tests has to be conducted separately; multiple comparison adjustments might be done via altered rejection regions.

2.1 Contrasts

Consider $\theta=\sum_{i=1}^a c_i\mu_i$ for some known c_i's with $\sum_{i=1}^a c_i=0$. One has $\hat{\theta}=\sum_{i=1}^a c_i\bar{Y}_i\sim N(\theta,\sigma_\theta^2)$, where $\sigma_\theta^2=\sigma^2\sum_{i=1}^a\frac{c_i^2}{n_i}$, and $\frac{\hat{\theta}-\theta_0}{s_p\sqrt{\sum_i c_i^2/n_i}}$ has df $(N-a)$ and noncentrality $\frac{\theta-\theta_0}{\sigma_\theta}$.

Tests concerning individual contrasts can be analyzed following the lines in §1.1; $\sum_i c_i=0$ is not used here, but general linear combinations of μ_i's are not of much interest in applications.

2.2 Overall F-Test

The F-test for $H_0 : \mu_1 = \cdots = \mu_a$ rejects H_0 when $\frac{\text{MSTr}}{\text{MSE}} > F_{1-\alpha,a-1,N-a}$, where $(a-1)\text{MSTr} = \text{SSTr} = \sum_{i=1}^{a} n_i(\bar{Y}_i - \bar{Y})^2$ for $\bar{Y} = \frac{1}{N}\sum_{i=1}^{a} n_i\bar{Y}_i$, and $\text{MSE} = s_p^2$ estimates σ^2. If σ^2 were known, one would be using $X^2 = \text{SSTr}/\sigma^2$ instead, rejecting H_0 when $X^2 > \chi^2_{1-\alpha,a-1}$.

X^2 follows a χ^2-distribution with df $(a-1)$ and noncentrality parameter $\phi = \sum_i n_i(\mu_i - \mu)^2/\sigma^2$, where $\mu = \frac{1}{N}\sum_{i=1}^{a} n_i\mu_i$, and $\frac{\text{MSTr}}{\text{MSE}}$ follows an F-distribution with df $(a-1, N-a)$ and the same noncentrality. With $a = 5$, $N = 35$, $\phi = 10$, and $\alpha = 0.05$, the power of the F-test is given by

```
1-pf(qf(.95,4,30),4,30,10)
```

Doubling the sample sizes in the setting, the power becomes

```
1-pf(qf(.95,4,65),4,65,20)
```

For the example in §1.2 with $|\mu_1 - \mu_2| = 0.8\sigma$ and $n_1 = n_2 = 15$, one has $a = 2$, $N = 30$, and $\phi = 15(2)(0.4)^2 = 4.8$. The power of the F-test here matches the power of the two-sided two-sample t-test in §1.2, as the two tests are equivalent.

```
1-pf(qf(.95,1,28),1,28,4.8)
1-pt(qt(.975,28),28,0.8*sqrt(15/2))+pt(-qt(.975,28),28,0.8*sqrt(15/2))
```

The noncentrality parameter $\phi = \sum_i n_i(\mu_i - \mu)^2/\sigma^2$ depends on the unknown σ^2 and μ_i's, for which one may use (conservative) ballpark estimates of σ^2 and hypothetical values of μ_i's. The following `R` function takes (n_i, μ_i) as inputs and returns $\sum_i n_i(\mu_i - \mu)^2$.

```
ncp=function(n,mu){mu0=sum(n*mu)/sum(n);sum(n*(mu-mu0)^2)}
ncp(5:9,c(3.9,4.1,4.2,4.3,4.5))
```

When n_i's are all equal and $\max|\mu_i - \mu_j| = \delta$, $\sum_{i=1}^{a}(\mu_i - \mu)^2 \geq \delta^2/2$, with the lower bound reached at $(\mu_1, \ldots, \mu_a) = \mu + (-\delta/2, 0, \ldots, 0, \delta/2)$.

3 Balanced Two-Way ANOVA

Consider $Y_{ijk} \sim N(\mu_{ij}, \sigma^2)$, $i = 1, \ldots, a$, $j = 1, \ldots, b$, $k = 1, \ldots, n$. One may write $\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$, with side conditions $\sum_i \alpha_i = \sum_j \beta_j = \sum_i(\alpha\beta)_{ij} = \sum_j(\alpha\beta)_{ij} = 0$. The ANOVA decomposition has SST = SSA+SSB+SSAB+SSE, where $\text{SST} = \sum_{i,j,k}(Y_{ijk} - \bar{Y}_{\cdots})^2$, $\text{SSA} = bn\sum_i(\bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdots})^2$, $\text{SSB} = an\sum_j(\bar{Y}_{\cdot j\cdot} - \bar{Y}_{\cdots})^2$, $\text{SSAB} = n\sum_{i,j}(\bar{Y}_{ij\cdot} - \bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdot j\cdot} + \bar{Y}_{\cdots})^2$, and $\text{SSE} = \sum_{i,j,k}(Y_{ijk} - \bar{Y}_{ij\cdot})^2$.

SSA/σ^2, SSB/σ^2, SSAB/σ^2, SSE/σ^2 are independent and all follow χ^2-distributions, with df's $(a-1)$, $(b-1)$, $(a-1)(b-1)$, $ab(n-1)$, in order, and noncentrality parameters $bn\sum_i \alpha_i^2/\sigma^2$, $an\sum_j \beta_j^2/\sigma^2$, $n\sum_{i,j}(\alpha\beta)_{ij}^2/\sigma^2$, 0.

When $\sum_{i,j}(\alpha\beta)_{ij}^2 > 0$, the two-way structure is largely artificial, and the setting effectively reduces to one-way ANOVA with ab treatment levels. The F-test for $H_0 : \sum_{i,j}(\alpha\beta)_{ij}^2 = 0$ rejects H_0 when $\frac{\text{SSAB}/(a-1)(b-1)}{\text{SSE}/ab(n-1)} > F_{1-\alpha,(a-1)(b-1),ab(n-1)}$, and the power analysis follows the lines of §2.2.

3.1 Additive Models

The two-way structure is effectively meaningful only when $\sum_{i,j}(\alpha\beta)^2_{ij} = 0$. Assuming an additive model $\mu_{ij} = \mu + \alpha_i + \beta_j$, one should combine SSAB and SSE, using $s_p^2 = \frac{\text{SSAB}+\text{SSE}}{abn-a-b+1}$ to estimate σ^2.

The F-test for $H_0 : \sum_i \alpha_i^2 = 0$ rejects H_0 when $\frac{\text{SSA}/(a-1)}{s_p^2} > F_{1-\alpha,a-1,abn-a-b+1}$, and the power analysis follows the lines of §2.2. The test for $H_0 : \sum_j \beta_j^2 = 0$ is similar.

For a one-way contrast $\theta = \sum_i c_i\alpha_i$ with $\sum_i c_i = 0$, $\hat\theta = \sum_i c_i\bar{Y}_{i\cdot\cdot} \sim N(\theta, \sigma_\theta^2)$, where $\sigma_\theta^2 = \sigma^2 \sum_i c_i^2/bn$, and $\frac{\hat\theta-\theta_0}{s_p\sqrt{\sum_i c_i^2/bn}}$ follows a t-distribution with df $(abn - a - b + 1)$ and noncentrality parameter $\frac{\theta-\theta_0}{\sigma_\theta}$. The power analysis follows the lines in §1.1. Contrasts of β_j's are similar.

3.2 Split-Plots

The standard design has a total of abn experimental units (EUs) evenly allocated to the ab cells of treatment combinations, and the index k has no meaning so can be arbitrarily permuted intra-cell.

Now suppose one can only assign the a levels of factor A to an blocks, but each block is divided into b subplots to receive the b levels of factor B. This leads to a split-plot design, with the an blocks as the EUs for factor A and the abn subplots as the EUs for factor B. One may write

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + g_{k(i)} + \epsilon_{ijk},$$

where $g_{k(i)} \sim N(0, \tau^2)$, independent of $\epsilon_{ijk} \sim N(0, \sigma^2)$.

The index k is now meaningful, nested under levels of factor A but crossed with levels of factor B, and $\sum_{i,j,k}(Y_{ijk} - \bar{Y}_{ij\cdot})^2$ involves both $g_{k(i)}$ and ϵ_{ijk}. Decomposing

$$\begin{aligned}\textstyle\sum_{i,j,k}(Y_{ijk} - \bar{Y}_{ij\cdot})^2 &= b\textstyle\sum_{i,k}(\bar{Y}_{i\cdot k} - \bar{Y}_{i\cdot\cdot})^2 + \sum_{i,j,k}(Y_{ijk} - \bar{Y}_{ij\cdot} - \bar{Y}_{i\cdot k} + \bar{Y}_{i\cdot\cdot})^2 \\ &= b\textstyle\sum_{i,k}(g_{k(i)} - \bar{g}_{\cdot(i)} + \bar\epsilon_{i\cdot k} - \bar\epsilon_{i\cdot\cdot})^2 + \sum_{i,j,k}(\epsilon_{ijk} - \bar\epsilon_{ij\cdot} - \bar\epsilon_{i\cdot k} + \bar\epsilon_{i\cdot\cdot})^2 \\ &= \text{SSBlk} + \text{SSE},\end{aligned}$$

one has SSE/σ^2 following a central χ^2-distribution with df $a(b-1)(n-1)$ and $\text{SSBlk}/(\sigma^2+b\tau^2)$ central χ^2 with df $a(n-1)$. $\text{SSAB} = n\sum_{i,j}((\alpha\beta)_{ij} + \bar\epsilon_{ij\cdot} - \bar\epsilon_{i\cdot\cdot} - \bar\epsilon_{\cdot j\cdot} - \bar\epsilon_{\cdot\cdot\cdot})^2$, so one may test $H_0 : \sum_{i,j}(\alpha\beta)^2_{ij} = 0$ using $\frac{\text{SSAB}/(a-1)(b-1)}{\text{SSE}/a(b-1)(n-1)}$, which has noncentrality $n\sum_{i,j}(\alpha\beta)^2_{ij}/\sigma^2$.

Tests for main effects make practical sense only in an additive model where $\sum_{i,j}(\alpha\beta)^2_{ij} = 0$, and one should pool resources and form $\text{SSE}^* = \text{SSE} + \text{SSAB}$ with df $(b-1)(an-1)$;

$$\text{SSE}^* = \textstyle\sum_{i,j,k}(Y_{ijk} - \bar{Y}_{i\cdot k} - \bar{Y}_{\cdot j\cdot} + \bar{Y}_{\cdot\cdot\cdot})^2 = \sum_{i,j,k}(\epsilon_{ijk} - \bar\epsilon_{i\cdot k} - \bar\epsilon_{\cdot j\cdot} + \bar\epsilon_{\cdot\cdot\cdot})^2.$$

$\text{SSA} = bn\sum_i(\alpha_i + \bar{g}_{\cdot(i)} - \bar{g}_{\cdot(\cdot)} + \bar\epsilon_{i\cdot\cdot} - \bar\epsilon_{\cdot\cdot\cdot})^2$, and one may test $H_0 : \sum_i \alpha_i^2 = 0$ using $\frac{\text{SSA}/(a-1)}{\text{SSBlk}/a(n-1)}$, which has noncentrality $bn\sum_i \alpha_i^2/(\sigma^2 + b\tau^2)$. Now $\text{SSB} = an\sum_j(\beta_j + \bar\epsilon_{\cdot j\cdot} - \bar\epsilon_{\cdot\cdot\cdot})^2$, and one may test $H_0 : \sum_j \beta_j^2 = 0$ using $\frac{\text{SSB}/(b-1)}{\text{SSE}^*/(b-1)(an-1)}$, which has noncentrality $an\sum_j \beta_j^2/\sigma^2$.

The power analysis of the above F-tests follows the lines of §2.2.

Contrasts of α_i's and β_j's can be analyzed following the routines used earlier, with the former involving $\sigma^2 + b\tau^2$ and the latter involving σ^2.

4 Proportions

Observing $Y_i \sim \text{Bin}(1,p)$, $i = 1,\ldots,n$, one is to test $H_0 : p \le p_0$ vs. $H_a : p > p_0$. One has $\mu = p$, $\sigma^2 = p(1-p) \le 0.5^2$, and $Y = \sum_{i=1}^n Y_i \sim \text{Bin}(n,p)$. One may use normal approximation to get started, then use `R` function `pbinom` to elicit via exact calculations. The normal approximation would set the rejection region as $Y > n(p_0 + z_{1-\alpha}\sqrt{p_0(1-p_0)/n})$.

Consider $p_0 = 0.3$, $\alpha = 0.05$, and one would like $p(0.4) \ge 0.8$; $\delta = \frac{0.4-0.3}{0.5} = 0.2$ to use in (2).

```
n=ceiling(((qnorm(.95)+qnorm(.8))/0.2)^2); n
c=ceiling(n*(.3+qnorm(.95)*sqrt(.3*(1-.3)/n))); c
1-pbinom(c,n,c(0.3,0.4))
```

So with $n = 155$ and rejection region $Y > 56$, one has $\alpha \le 0.042$ and $1 - \beta \ge 0.816$ for $p \ge 0.4$. In general, one may need to fiddle a bit around the results based on normal approximation.

Two-sided and/or two-sample situations could be tricky, but one should be able to get around using normal approximation, and possibly perform exact power calculations via `pbinom` using hypothetical true values of p. The sample sizes needed here are usually large, so normal approximation should be safe unless p gets too close to 0 or 1.