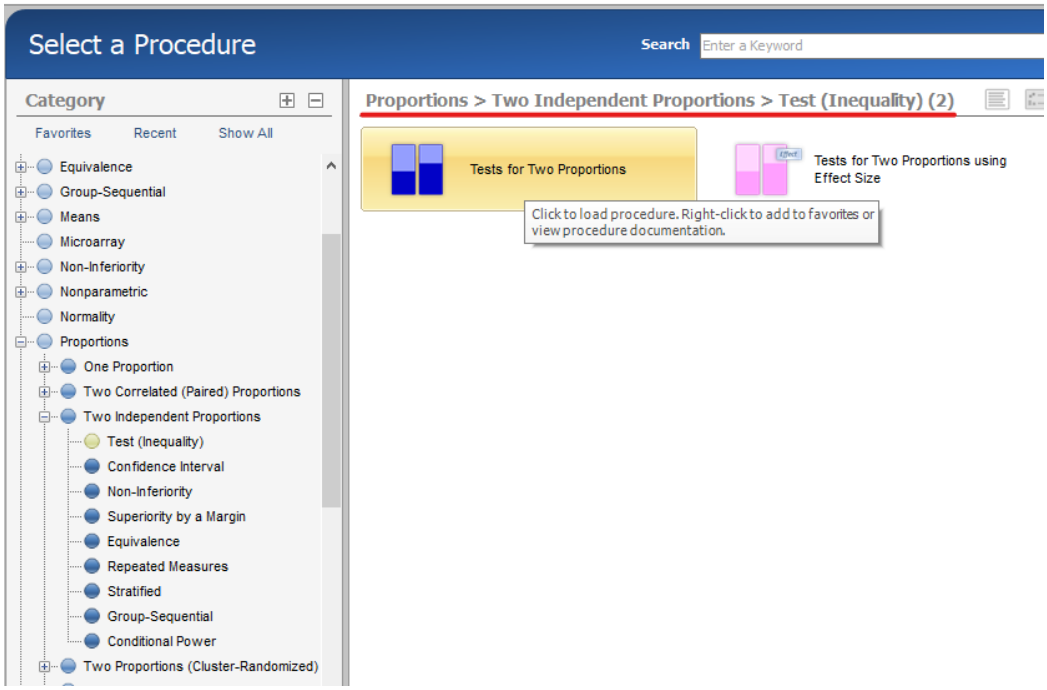
医学统计样本量估计及PASS软件应用

- **介绍**
  - **应用场景**
    - 确定研究所需的样本含量
    - 已获取样本数据，判断是否满足研究需要
  - **样本量估计和判断的依据标准**
    - 检验水平——显著性水平α
    - 检验能效——统计检验力power（1-β）
    - α和β具体含义
      - α为拒绝了真实的原假设的概率，即拒真概率，称为犯第一类错误的概率；β为接受了错误的原假设的概率，即纳伪概率，称为犯第二类错误的概率。一般我们要保证在犯第一类错误的概率一定的情况下，使犯第二类错误的概率尽可能小。

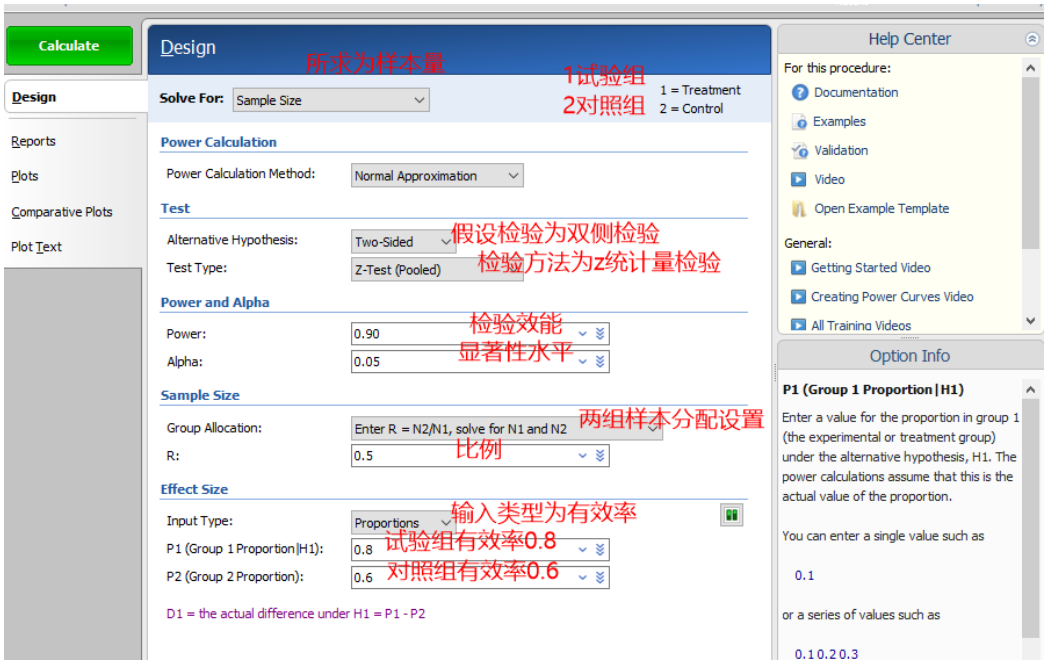
		统计检验结果	
		拒绝H0	接受H0
真实情况	H0成立	α	1-α
	H0不成立	1-β	β

- **案例讲解**
  - **案例1——完全随机设计、两样本率比较时样本含量估计方法**
    - 例：某医生拟对两种抗菌药物（其中一种为对照药）治疗某种感染性疾病的药物疗效进行研究。经预试验，发现试验药有效率为80%，对照组有效率为60%，现两组药物样本比例按照2:1比例平行设计、欲在α=0.05（双侧检验），检验能效（1-β）=0.9时发现两组间有差异，①问正式临床试验各组需要多大的样本容量？②当样本容量为200时，是否能满足本研究需要？

- 解释：目的是进行两样本率差异性检验时对所需要的样本含量进行估计。根据题目可知五个参数分别为：①PT（试验组有效率）=80%；②PC（对照组有效率）=60%；③双侧检验α（显著性水平）=0.05；④检验效能（1-β）=0.9；⑤试验组样本与对照组样本之比为2:1。
- 操作步骤：
  - 第一步



  - 第二步



  - 结果解读

Tests for Two Proportions

Numeric Results for Testing Two Proportions using the Z-Test with Pooled Variance
H0: P1 - P2 = 0. H1: P1 - P2 = D1 ≠ 0.

Target Power	Actual Power*	N1	N2	N	Target R	Actual R	P1	P2	Diff D1	Alpha
0.90	0.90011	159	80	239	0.50	0.50	0.8000	0.6000	0.2000	0.0500

* Power was computed using the normal approximation method.

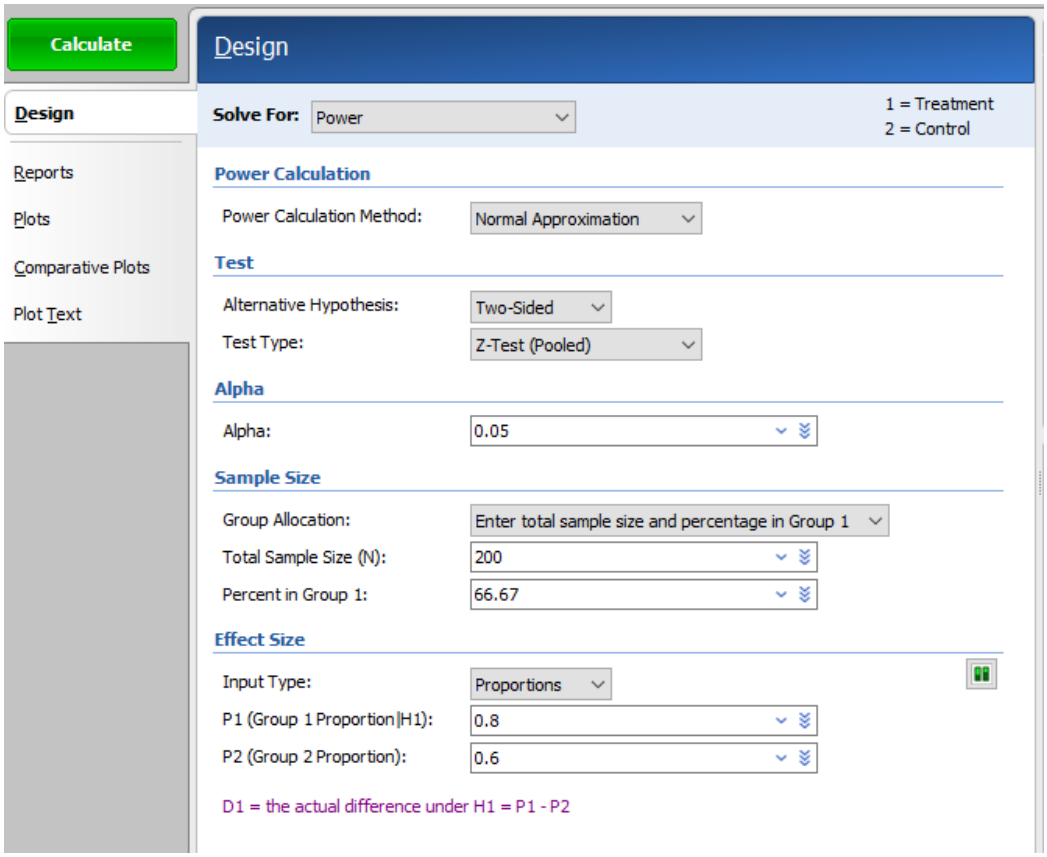
References

Chow, S.C., Shao, J., and Wang, H. 2008. Sample Size Calculations in Clinical Research, Second Edition. Chapman & Hall/CRC. Boca Raton, Florida.
D'Agostino, R.B., Chase, W., and Belanger, A. 1988. 'The Appropriateness of Some Common Procedures for Testing the Equality of Two Independent Binomial Populations', The American Statistician, August 1988, Volume 42 Number 3, pages 198-202.
Fleiss, J. L., Levin, B., and Paik, M.C. 2003. Statistical Methods for Rates and Proportions. Third Edition. John Wiley & Sons. New York.
Lachin, John M. 2000. Biostatistical Methods. John Wiley & Sons. New York.
Machin, D., Campbell, M., Fayers, P., and Pinol, A. 1997. Sample Size Tables for Clinical Studies, 2nd Edition. Blackwell Science. Malden, Mass.
Ryan, Thomas P. 2013. Sample Size Determination and Power. John Wiley & Sons. Hoboken, New Jersey.

Report Definitions

Target Power is the desired power value (or values) entered in the procedure. Power is the probability of rejecting a false null hypothesis.
Actual Power is the power obtained in this scenario. Because N1 and N2 are discrete, this value is often (slightly) larger than the target power.
N1 and N2 are the number of items sampled from each population.
N is the total sample size, N1 + N2.
Target R is the desired ratio (or ratios) of R entered in the procedure. R is the ratio of N2 to N1, so that N2 = R × N1.
Actual R is the value for R obtained in this scenario. Because N1 and N2 are discrete, this value is sometimes slightly different than the target R.
P1 is the proportion for Group 1 at which power and sample size calculations are made. This is the treatment or experimental group.

- 第四步



- 结果解读

Tests for Two Proportions

Numeric Results for Testing Two Proportions using the Z-Test with Pooled Variance
H0: P1 - P2 = 0. H1: P1 - P2 = D1 ≠ 0.

Power*	N1	N2	N	Target %N1	Actual %N1	P1	P2	Diff D1	Alpha
0.84226	133	67	200	66.7	66.5	0.8000	0.6000	0.2000	0.0500

* Power was computed using the normal approximation method.

References

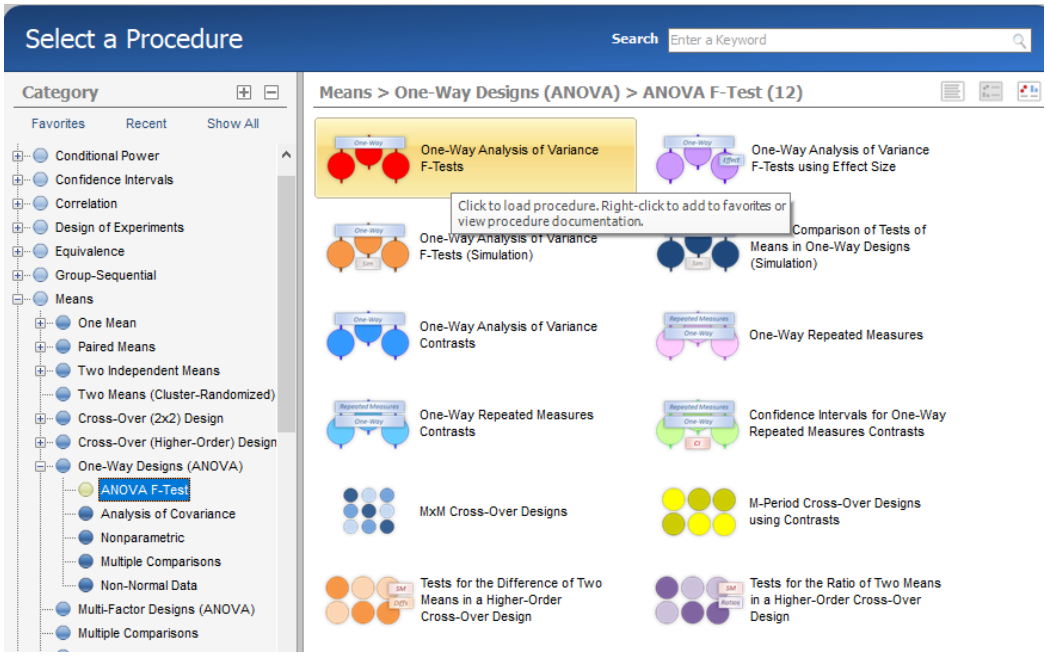
Chow, S.C., Shao, J., and Wang, H. 2008. Sample Size Calculations in Clinical Research, Second Edition. Chapman & Hall/CRC. Boca Raton, Florida.
D'Agostino, R.B., Chase, W., and Belanger, A. 1988. 'The Appropriateness of Some Common Procedures for Testing the Equality of Two Independent Binomial Populations', The American Statistician, August 1988, Volume 42 Number 3, pages 198-202.
Fleiss, J. L., Levin, B., and Paik, M.C. 2003. Statistical Methods for Rates and Proportions. Third Edition. John Wiley & Sons. New York.
Lachin, John M. 2000. Biostatistical Methods. John Wiley & Sons. New York.
Machin, D., Campbell, M., Fayers, P., and Pinol, A. 1997. Sample Size Tables for Clinical Studies, 2nd Edition. Blackwell Science. Malden, Mass.
Ryan, Thomas P. 2013. Sample Size Determination and Power. John Wiley & Sons. Hoboken, New Jersey.

Report Definitions

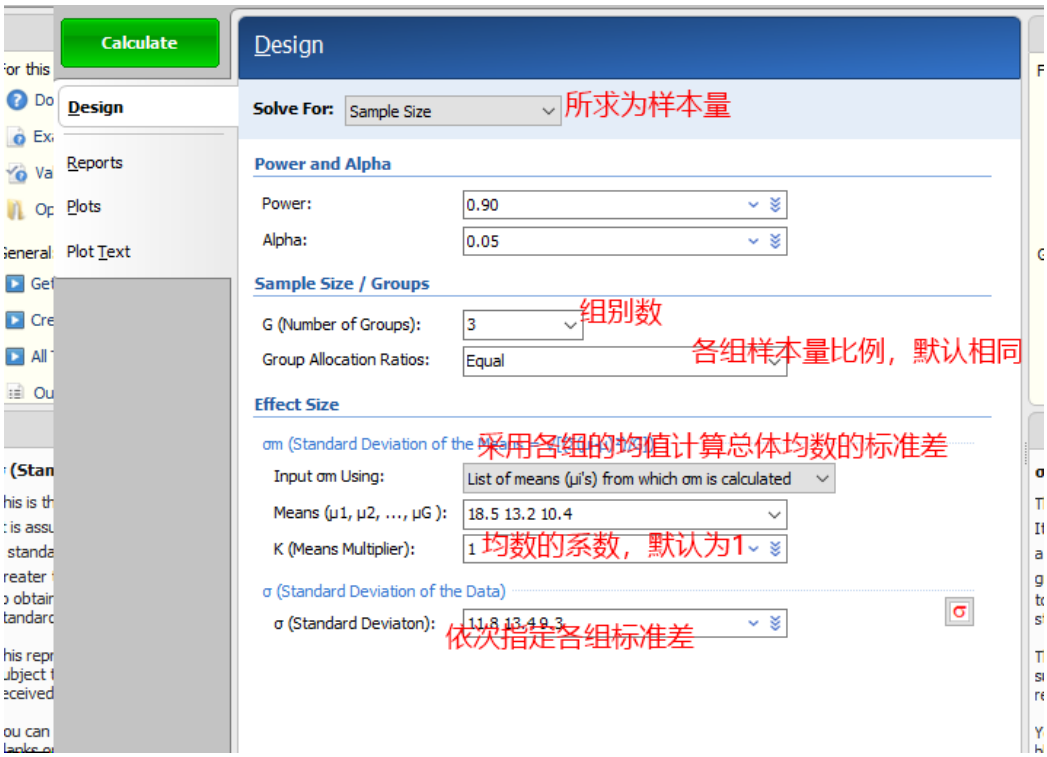
Power is the probability of rejecting a false null hypothesis.
N1 and N2 are the number of items sampled from each population.
N is the total sample size, N1 + N2.
Target %N1 is the desired percent (or percents) of the total sample size to be allocated to Group 1, as entered in the procedure.
Actual %N1 is the percent in N1 obtained. Because N1 and N2 are discrete, this value is sometimes slightly different than the target %N1.
P1 is the proportion for Group 1 at which power and sample size calculations are made. This is the treatment or experimental group.
P2 is the proportion for Group 2. This is the standard, reference, or control group.
D1 is the difference P1 - P2 assumed for power and sample size calculations.

- **案例二——完全随机设计多样本均数比较样本量估计**
  - 当研究完全随机设计的两均数比较的样本量估计时，其研究处理的因素只有一个且仅为两个水平，这时采用两独立样本T检验的方法计算样本含量。但在医学研究中，经常遇到处理因素有k个水平（k≥30）的情况，如果不同药物剂量间疗效的比较、临床不同治疗方案的比较等，这时应采用单因素方差分析ANOVA检验来计算样本含量。
  - 例：某医生观察三种降压药的降压效果，经预实验测得各药物治疗后血压下降的均数分别为18.5mmHg、13.2mmHg和10.4mmHg，标准差分别为11.8mmHg、13.4mmHg和9.3mmHg。若要求α=0.05（双侧检验），检验效能1-β=0.9时，各组主要多少病人进行临床试验？（血压单位毫米汞柱mmHg）
  - 解析：本例是个完全随机设计的研究，其实验因素具有三水平：三种不同的降压药，主要结局指标是血压下降水平，是连续型变量。目的是进行多样本均数比较的样本含量估计。根据题目我们知道了四组参数：①各组的均数；②各组的标准差；③α=0.05（双侧检验）；④检验效能1-β=0.9
  - 操作步骤：

- 单因素方差分析(One-Way ANOVA)是一种用于比较三个或三个以上组别间均值差异的统计方法
- 第一步



- 第二步



- 结果解读

One-Way Analysis of Variance F-Tests

Numeric Results
Means: 18.5 13.2 10.4

Power	Average n	G	Total N	K	Std Dev of Means σm	Standard Deviation σ	Effect Size	Alpha
0.9059	34.00	3	102	1.00	3.36	9.30	0.3612	0.0500
0.9053	54.00	3	162	1.00	3.36	11.80	0.2847	0.0500
0.9038	69.00	3	207	1.00	3.36	13.40	0.2507	0.0500

References
Desu, M. M. and Raghavarao, D. 1990. Sample Size Methodology. Academic Press. New York.
Fleiss, Joseph L. 1986. The Design and Analysis of Clinical Experiments. John Wiley & Sons. New York.
Kirk, Roger E. 1982. Experimental Design: Procedures for the Behavioral Sciences. Brooks/Cole. Pacific Grove, California.

Report Definitions
Power is the probability of rejecting a false null hypothesis. It should be close to one.
n is the average group sample size.
G is the number of groups.
Total N is the total sample size of all groups combined.
K is the group means multiplier.
σm is the standard deviation of the group means under the alternative hypothesis.
σ is the within group standard deviation.
The Effect Size is the ratio of σm and σ.
Alpha is the probability of rejecting a true null hypothesis. It should be small.

Summary Statements
In a one-way ANOVA study, sample sizes of 34, 34, and 34 are obtained from the 3 groups whose means are to be compared. The total sample of 102 subjects achieves 91% power to detect differences among the means versus the alternative of equal means using an F test with a 0.0500 significance level. The size of the variation in the means is represented by their standard deviation which is 3.36. The common standard deviation within a group is assumed to be 9.30.

- 各组标准差不一样，故对应三种样本量结果。