Sample size calculation

最小观察例数
两个角度：

- from confidence interval（多用于单样本）
- from hypothesis testing（多用于多样本or样本比较）

from CI

basic

Recall: two-sided CI for one-sample

Population Distribution	Sample Size	Population Variance	95% Confidence Interval
Binomial	Large	-	$\hat{p} \pm 1.96\sqrt{\hat{p}(1-\hat{p})/n}$
	Small	-	Exact methods

Population Distribution	Sample Size	Population Variance	95% Confidence Interval
Normal	Any	σ^2 known	$\bar{X} \pm 1.96\sigma/\sqrt{n}$
	Any	σ^2 unknown, use s^2	$\bar{X} \pm t_{0.025,n-1}s/\sqrt{n}$
Not Normal/ Unknown	Large	σ^2 known	$\bar{X} \pm 1.96\sigma/\sqrt{n}$
	Large	σ^2 unknown, use s^2	$\bar{X} \pm 1.96s/\sqrt{n}$
	Small	Any	Non-parametric methods

margin of error ?

计算一个样本的双边CI
CI常规来说和n都是有关系的

it is desirable to have ***narrow*** CIs

- smaller sample size → wider CIs 只能改变n
- higher confidence → wider CIs 预先设定这是啥
- larger variability → wider CIs 无法控制
  注意：n和CI并非完全是正比关系（而且本来就是反向关系

In studies concerned with detecting an effect

- If an effect deemed to be clinically or biologically important exists, then there is a high chance of it being detected, i.e. that the analysis will be statistically significant.
- If the sample is too small, then even if large differences are observed, it will be impossible to show that these are due to anything more than sampling variation.

如果样本上太小，就算生物学意义上有显著性，但也观察不到statistically significant

> statistically significance: （确实看到了有效的结果）但因为只出现了一次，在统计学上可能性太小，鉴定为随机出现的

样本量不够大导致了 ***不稳定***

example

估计单个平均值

1.3 Example 1 - estimating a single mean

An investigator wants to estimate the mean systolic blood pressure in children with congenital heart disease. How many children should be enrolled in the study? The investigator plans on using a 95% confidence interval (so $Z_{\alpha/2}$ =1.96) and wants a **margin of error (MOE) of 5 units**.

The standard deviation of systolic blood pressure is unknown, but the investigators conduct a literature search and find that the standard deviation of systolic blood pressures in children with other cardiac defects is between 15 and 20.

To estimate the sample size, we consider the larger standard deviation in order to obtain the most conservative (largest) sample size.

sd: unknown
可以15、20都算一遍，选算出的更大的n

1.3 Example 1 - estimating a single mean

- CI: $\bar{X} \pm \boxed{Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}$ MOE 容许误差
- Sample size:

$$\text{MOE} = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 5$$ ←不是两倍

$$n = \left(Z_{\alpha/2} \cdot \sigma/\text{MOE}\right)^2 = (1.96 * 20/5)^2 = 61.5$$

Sample size = 62 取整

62指的是 *有效* 样本（可能有人没配合

example 2

1.3 Example 2 - Estimating a single proportion

- Scenario: The prevalence of dysfunctional breathing amongst asthma patients being treated in general practice is to be assessed using a postal questionnaire survey (Thomas *et al.* 2001).
- Required information:
  - Primary outcome variable = presence/absence of dysfunctional breathing
  - 'Best guess' of expected percentage (proportion) = 30% (0.30)
  - Desired width of 95% confidence interval = 10% (i.e. +/- 5%, or 25% to 35%)

0-1 outcome（presence/absence）
二项分布（且CLT） 不是正态分布
所以用z不是t 而且t分布一般针对连续型变量，此处结果是二分的

1.3 Example 2 - Estimating a single proportion

- CI: $\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$

$(\hat{p} \ \pm MOE)$ Point estimate ± Margin of Error

- Formula for sample size for estimation of a proportion is

$$n = \hat{p}(1-\hat{p})(Z_{\alpha/2}/\mathrm{MOE})^2$$

  - where n = the required sample size
  - $\hat{p}$ = the proportion estimation - here 0.30
  - MOE = margin of error here 0.05 上文中10%的一半

phat不能太偏而且np需要检查，这样才能估计

1.3 Example 2： Estimating a single proportion

- Here we have: $n = 1.96^2 * 0.30 * (0.70)/0.05\text{^}2 = 322.7$
- "A sample of 323 patients with asthma will be required to obtain a 95% confidence interval of +/- 5% around a prevalence estimate of 30%. To allow for an expected 70% response rate to the questionnaire, a total of 480 questionnaires will be delivered."
- Note: The formula presented below is based on 'normal approximation methods', and, should not be applied when estimating percentages which are close to 0% or 100%. In these circumstances 'exact methods' should be used.

考虑到有些样本不有效
如果phat很偏的话就不能使用（比如phat=0.03 eg.罕见病发病率，这样就不能用正态分布估计，需要使用exact method（用二项分布算））
从n=5算到n=1000000000一直到moe符合要求ry（就不用公式推）

from hypothesis testing

Power（probability）

前提：H_a=True

power=probility(reject H_0 | H_a) ←正确地拒绝了H0（真阳性
eg抛硬币-1.所有都达标：是好人 2. 五项有三项达标：是好人
捞到好人的概率更大/后者的敏感性（即power更好
ideal test：在α小的同时 尽可能使power大

power=probility(rejection region | H_a) ⇐⇒知道H_a下的sampling dist

面积←→概率

*替代终点*不一定要找检验统计量

2.1 Power (1) – definition

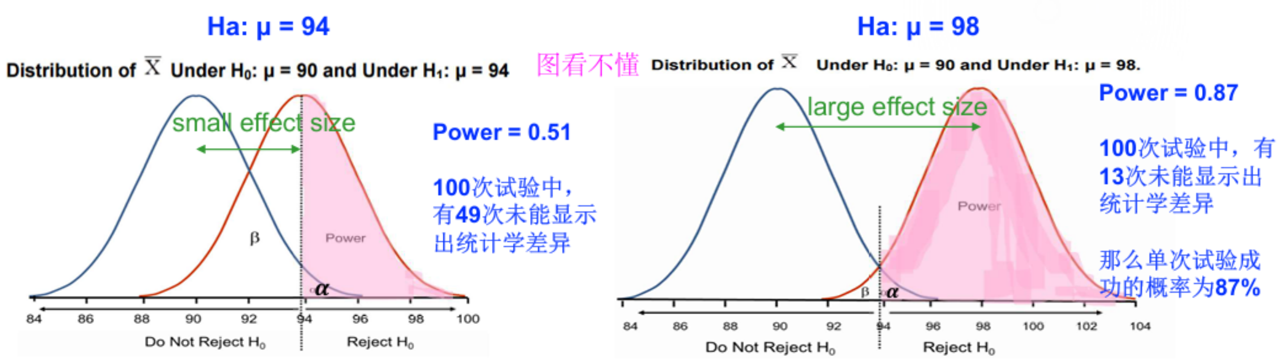
13

单边

- Suppose we want to test the hypotheses at a α=0.05: H0: μ = 90 versus Ha: μ > 90. To test the hypotheses, suppose we select a sample of size n=100.

效应量

- Effect size: We can think about our effect size as the importance of a certain effect. The larger the effect size, the more easily it can be seen by just looking.



α等大

粉色的是1-β=power

同样的检测方法，如果效应足够明显（μ很大），就容易检测出来。large effect size：98-90=8 > small effect size 94-90=4

power的含义（见图）作图确实有效应，但监测不出来（很有可能）

example

需要：

- sampling dist(under Ha)

- reject region

2.1 Power (2) – Example*

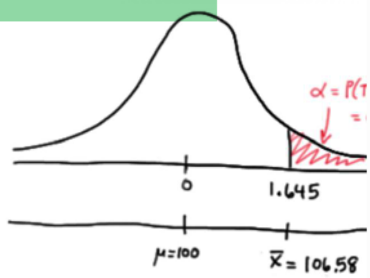
14

Let X denote the IQ of a randomly selected adult American. Assume, a bit unrealistically, that X is normally distributed with unknown mean μ and standard deviation 16.

Take a random sample of $n = 16$ students, after setting the probability of committing a Type I error at $\alpha = 0.05$, we can test the null hypothesis $H_0 : \mu = 100$ against the alternative hypothesis that $H_A : \mu > 100$.

What is the power of the hypothesis test if the true population mean were $\mu = 108$?

Answer: setting $\alpha = 0.05$, implies that we should reject the null hypothesis when the test statistic $Z \geq 1.645$, or equivalently, when the observed sample mean is 106.58 or greater:



$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad \Rightarrow \bar{X} = \mu + Z\frac{\sigma}{\sqrt{n}} \qquad \bar{X} = 100 + 1.645\left(\frac{16}{\sqrt{16}}\right) = 106.58$$

≥106.58 reject

（题中sd=n 不要搞混）

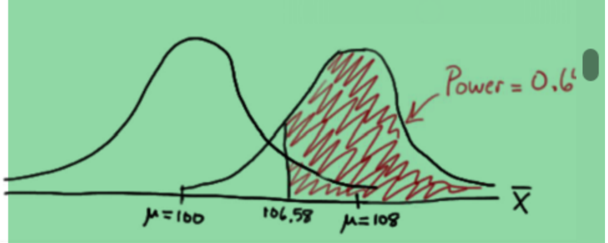
15 / 28

2.1 Power (2) – Example*

15

$$\text{Power} = P(\bar{X} \geq 106.58 \text{ when } \mu = 108) = P\left(Z \geq \frac{106.58 - 108}{\frac{16}{\sqrt{16}}}\right) = P(Z \geq -0.36) = 1 - P(Z < -0.36)$$

We have determined that we have (only) a 64.06% chance of rejecting the null hypothesis H0:μ=100 in favor of the alternative hypothesis Ha:μ>100 if the true unknown population mean is in reality μ=108.



几何意义
（因为sampling dist是正态分布所以这么算了）
可以看到前一张图中Ha的μ已经给定了

2.1 Power (3) – calculation*

16

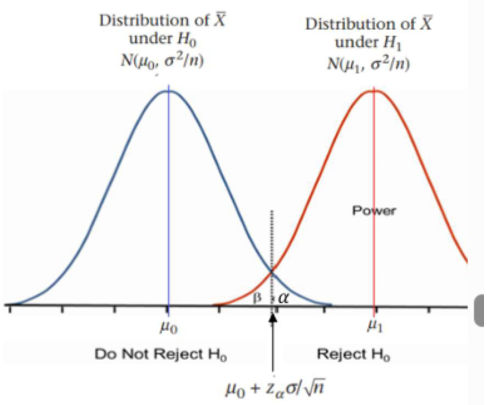
H0: $\mu = \mu_0$ vs. Ha: $\mu > \mu_0$

Power = P(reject H0|H0 false)

$$= P(Z > z_{1-\alpha} | \mu = \mu_1)$$

$$= P\left(\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} > z_{1-\alpha} \middle| \mu = \mu_1\right)$$

$$= P(\bar{x} > \mu_0 + z_{1-\alpha}\sigma/\sqrt{n} | \mu = \mu_1)$$



Under Ha, $\bar{X} \sim N(\mu_1, \sigma^2/n)$. Hence, after standardization,

$$\text{Power} = 1 - \Phi\left[\frac{\mu_0 + \frac{z_{1-\alpha}\sigma}{\sqrt{n}} - \mu_1}{\frac{\sigma}{\sqrt{n}}}\right] = 1 - \Phi[z_{1-\alpha} + \frac{\mu_0-\mu_1}{\sigma}\sqrt{n}]$$

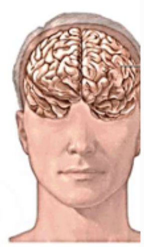
↑ σ已知
如果σ未知：t分布 在H0下是0为中心的t分布，但Ha下不是0为中心，很难算
从上图中可得**想要达到一定的power**需要调节n（前面那个跟抛硬币一样的概率，差异确实存在，但要想检测到比较小的差距就需要增加样本量）

example-续

2.2 Calculating sample size

17

Let X denote the IQ of a randomly selected adult American. Assume, a bit unrealistically again, that X is normally distributed with unknown mean μ and (a strangely known) standard deviation of 16. This time, instead of taking a random sample of $n = 16$ students, let's increase the sample size to $n = 64$. And, while setting the probability of committing a Type I error to $\alpha = 0.05$, test the null hypothesis $H_0 : \mu = 100$ against the alternative hypothesis that $H_A : \mu > 100$.

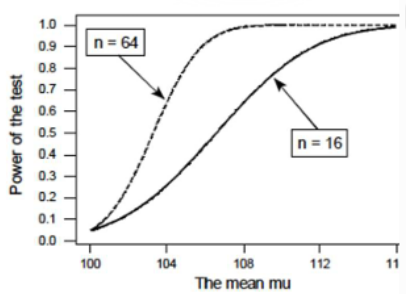


What is the power of the hypothesis test when $\mu = 108$, $\mu = 112$, and $\mu = 116$?

Answer: we can calculate the power of testing $H0: \mu = 100$ against $Ha: \mu > 100$ for 2 sample sizes (n=16 and n=64) and for 3 possible values of the mean (μ=108, μ=112, and μ=116). Here's a summary of our power calculations:

POWER	$K(108)$	$K(112)$	$K(116)$
$n = 16$	0.6406	0.9131	0.9909
$n = 64$	0.9907	0.9999...	0.999999...

pow ↑ ↑ ↑



清华大学统计学研究中心

pow都不断接近0.9999..
power curve μ-power

summary

2.2 Calculating sample size

18

- Recall Power = $1 - \Phi\left[\frac{\mu_0 + \frac{z_{1-\alpha}\sigma}{\sqrt{n}} - \mu_1}{\frac{\sigma}{\sqrt{n}}}\right] = 1 - \Phi[z_{1-\alpha} + \frac{\mu_0 - \mu_1}{\sigma}\sqrt{n}]$ 1-sided test
- The sample size computations depend on the level of significance α, the desired power of the test (equivalent to $1-\beta$), the variability of the outcome, and the effect size.

？？？？？？？？

One-sided test: $n = \frac{(Z_{1-\alpha}+Z_{1-\beta})^2}{(\mu_1 - \mu_0)^2/\sigma^2}$ 实际差距越大 所需的样本量就越小

if σ^2 is unknown, replace σ^2 by S^2

Two-sided test: $n = \frac{(Z_{1-\alpha/2}+Z_{1-\beta})^2}{(\mu_1 - \mu_0)^2/\sigma^2}$

other scenarios

one sample proportion

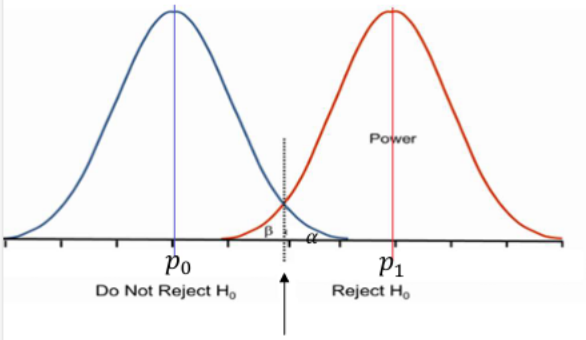
3.1 Sample size for one sample proportion – one-sided

22

(assuming large sample)

Under H0: $\hat{p} \sim N(p_0, \frac{p_0(1-p_0)}{n})$

Under Ha: $\hat{p} \sim N(p_1, \frac{p_1(1-p_1)}{n})$



$H_0: p = p_0$ vs. $H_a: p <$ or $> p_0$

(we believe true $p = p_1$)

$$n = \frac{\left(Z_{1-\alpha}\sqrt{p_0(1-p_0)} + Z_{1-\beta}\sqrt{p_1(1-p_1)}\right)^2}{(p_1 - p_0)^2}$$

$$^*1-\beta = \Phi\left(\sqrt{\frac{p_0(1-p_0)}{p_1(1-p_1)}}\left(\frac{|p_1-p_0|\sqrt{n}}{\sqrt{p_0(1-p_0)}} - Z_{1-\alpha}\right)\right)$$

通过rejection region得到power

Φ是啥意思

Φ (x)是X~N (0，1)时X的 分布函数 。 分布函数（英文Cumulative Distribution Function, 简称CDF）

清华大学统计学研究中心

- large sample
- one sample
- 中心位置不同，方差也不同（可能形状也发生变化）
- 双边检验（?
- **we believe true p=p1**←值已经给定
  实际上p1只要大于p0就能接受，但在计算的时候需要给一个值，不能模糊地说大于p0（。）

效应不足可以通过增加样本量弥补

乐观估计的效应过大导致样本量不够就很不好

two-population mean

比较差异-作差

$$H_0: \mu_A - \mu_B = 0$$
$$H_1: \mu_A - \mu_B \neq 0$$

The ratio between the sample sizes of the two groups is $\kappa = \frac{n_A}{n_B}$

看不懂

$$n_A = \kappa n_B \text{ and } n_B = \left(1 + \frac{1}{\kappa}\right)\left(\sigma \frac{z_{1-\alpha/2} + z_{1-\beta}}{\mu_A - \mu_B}\right)^2$$

$$1 - \beta = \Phi\left(z - z_{1-\alpha/2}\right) + \Phi\left(-z - z_{1-\alpha/2}\right) \quad , \quad z = \frac{\mu_A - \mu_B}{\sigma\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}}$$

two-population proportion

同样也是作差 然后利用K

3.3 Sample size for two-population proportions

The ratio between the sample sizes of the two groups is $\kappa = \frac{n_A}{n_B}$

$$H_0: p_A - p_B = 0$$
$$H_1: p_A - p_B \neq 0$$

$$n_A = \kappa n_B \text{ and } n_B = \left(\frac{p_A(1-p_A)}{\kappa} + p_B(1-p_B)\right)\left(\frac{z_{1-\alpha/2} + z_{1-\beta}}{p_A - p_B}\right)^2$$

$$1 - \beta = \Phi\left(z - z_{1-\alpha/2}\right) + \Phi\left(-z - z_{1-\alpha/2}\right) \quad , \quad z = \frac{p_A - p_B}{\sqrt{\frac{p_A(1-p_A)}{n_A} + \frac{p_B(1-p_B)}{n_B}}}$$

以上两个都是正态的比较容易

paired sample

作差→one-sample case

summary

数据类型不同-test 方法不同-设计方案不同-样本量的估计方法不同
估算的样本量是最少需要的量，考虑可能出现的各种情况，在计划样本量时候应增加若干样本量例数

（e.g., +20%）
多指标都计算取最大值or只对最主要指标进行样本量估计

实际情况：

- 不能违背伦理
- 多组设计时不一定各组样本量相等（20实验组 10控制组）

Sign test, Wilcoxon test, and Normality test

非参数检验
对所有数据都适用，无论该数据是否能使用参数检验
非参数检验的power更大，敏感性更好

non-parametric hypothesis test

参数：p μ p1-p2 μ1-μ2
这些参数 **的分布是确定的**，data population dist已知
可以据此写出检验统计量T的特别明确的sampling dist（配合CLTry）
非参数检验：
（eg 对整体μ1、μ2的关系感兴趣，本身数据属于什么分布不知道也不感兴趣或者过于奇怪，又想比较population mean
也可以构建检验统计量T，也有分布（但可能比较复杂
（eg 对数据的分布感兴趣，数据是否服从正态分布？（不管参数是多少
使用条件：

- 达不到分布的要求，不能说是XX分布
  - population dist
    =Normal
    - σ→normal
    - s→t dist
  - *≠normal*
    - sample size large↑
    - *small*
  - ↑条件不满足的时候这些方法不能用
- 数据本身就是关于rank ranked data
  - 很明显不是正态分布 1 2 3 ...
  - *parametric tests require stonger scale than rank*
- 不关心参数
  - eg *is the sample random?* 只关心随机性
  - *is the sample from population following a normal distribution?* 只关心分布类型
- 言而总之应用很广泛

sign test符号检验

eg：检验药效

- Compare the effects of two soporific drugs

Subject	Drug 1	Drug 2	Diff (2-1)
1	1.9	0.7	−1.2
2	−1.6	0.8	2.4
3	−0.2	1.1	1.3
4	−1.2	0.1	1.3
5	−0.1	−0.1	0.0
6	3.4	4.4	1.0
7	3.7	5.5	1.8
8	0.8	1.6	0.8
9	0.0	4.6	4.6
10	2.0	3.4	1.4

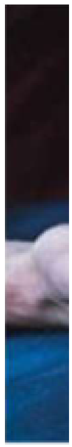


Table: Hours of extra sleep on drugs 1 and 2, differences, signs and ranks of sleep study data

注意看标题这是睡眠增量不是总时长

- paired comparison
  - paired t test（前提：normal）paired=dependent 还有two sample t test
  - 此处对数据分布类型不感兴趣
- rv
  - 用Xi表示drug 1 Yi表示drug 2，Xi independent; Yi independent *Xi, Yi are dependent-- cannot use two sample t test*
  - Xi之间的差异是人导致的，X和Y之间的差异是药导致的，前者的平均差异应该为0
  - Dfference Zi=Yi-Xi Zi是一个rv 药有效 即 E(Z)≠0
- test
  - t-test?
    - t tests for the [means] of [continuous] data
    - normal dist (接近正态但不需要完全正)

1.1.1 Examples of when T test Goes Wrong

- Extreme outliers
  - Example: t test comparing two sets of measurements
    - Sample 1: 1 2 3 4 5 6 7 8 9 10
    - Sample 2: 7 8 9 10 11 12 13 14 15 16 17 18 19 20
    - Sample average: 5.5 and 13.5
    - Sample variance: 9.2 and 17.5
    - t test: $t = -5.4349$; P-value: $p = 0.000019$
  - Example: t test comparing two sets of measurements
    - Sample 1: 1 2 3 4 5 6 7 8 9 10
    - Sample 2: 7 8 9 10 11 12 13 14 15 16 17 18 19 20 200
    - Sample average: 5.5 and 25.9
    - Sample variance: 9.2 and 2335
    - t test: $t = -1.6329$; P-value $p = 0.12$

  - 根据后者p值给出的结论：没有区别。这一结论很显然不靠谱
- **Sign test** 类似t test 非参数版 **可以用于检验paired data；也可以检验 *中位数***

calculate

把差异正负转化成0和1

1.2 Sign Test

- Create a sign for each $Z_i, i = 1, \cdots, n$ $\quad S_i = \begin{cases} 1 & Z_i > 0 \\ -1 & Z_i < 0 \\ 0 & Z_i = 0 \end{cases}$
- H_0: The mean effects of two drugs are the same, or equivalently

Under H_0, what is the properties of S_i?

多选题 1分

已完成

10

Create a sign for each $Z_i, i = 1, \cdots, n$ Zi 观测值

$$S_i = \begin{cases} 1 & Z_i > 0 \\ -1 & Z_i < 0 \\ 0 & Z_i = 0 \end{cases}$$

H_0: The mean effects of two drugs are the same, or equivalently.

Under H_0, what are the properties of S_i?

A The distribution of Z_i's is symmetric about its mean

【零不考虑】 P(Si=0)=0

B For any given subject, the difference in hours of sleep equally likely to be positive or negative, $P(S_i = 1) = P(S_i = -1) = 0.5$

C The number of positive Z_i's, n_+, can be regarded as binomial distribution. 公平的硬币

D Let $n_0 = n_+ + n_-$, then $n_+ \sim B(n_0, \theta), H_0\!: \theta = 0.5 \leftrightarrow H_1\!: \theta \neq 0.5$

说点什么

清华大学统计学研究中心

差异为0的情况并不完全存在 （=0的数据剔除不用（？））

P(Si=1)=P(Si=-1)=0.5

n+≈n/2 不能偏得太离谱

The rejection region at level of significance α takes the form

$$D = \{\boldsymbol{X} = (X_1, \cdots, X_n) : n_+ \geq c \text{ or } n_+ \leq d\}$$

确定c d的值

The constants c, d can be determined by

$$\text{二项分布的概率} \sum_{i=c}^{n_0} C_{n_0}^i 0.5^{n_0} \leq \frac{\alpha}{2}, \quad d = n_0 - c$$

c+d=n0，n0是由α确定的，是非常精确的exact test

- Compute the P-value of the test. Let $x_0 = \min\{n_+, n_0 - n^+\}$ ？？

$$p = \sum_{i=0}^{x_0} C_{n_0}^i 0.5^{n_0} + \sum_{i=n_0-x_0}^{n_0} C_{n_0}^i 0.5^{n_0}$$

(If n_0 is even and $n_+ = n_0/2$, then define $p = 1$.)

- Give a level of significance α, we reject H_0 if $p < \alpha$.

通过Z到S的转化，发现这是一个二项分布，通过二项分布的特点进行exact test

如果n0正好是一半那就非常完美，p=1

1.2 Sign Test – back to example

- Drug example
  - $n_+ = 8, n_0 = 9$
  - (Exact) P-value (probability of observing 0,1,8,9 positives): $p = 0.0195 < 0.05$
  - Reject H_0 at level of significance $\alpha = 0.05$

R will help you

sign test- test median

所有的数对中位数作差，有一半正数一半负数

真就符号检验

1.3 Sign Test – test median

14

- Sign test can be used to test about the median of a population

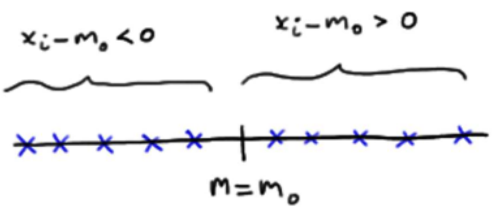
$$H_0: m = m_0$$

$$H_A: m > m_0 \text{ or } H_A: m < m_0 \text{ or } H_A: m \neq m_0$$

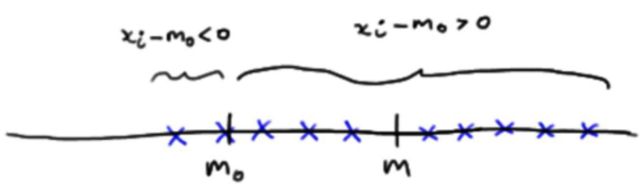
What is m ?

What is m_0?

If the **null hypothesis is true**, then we should expect about half of the $x_i - m_0$ quantities obtained to be positive and half to be negative:



If instead, $m > m_0$, then we should expect **more** than half of the $x_i - m_0$ quantities obtained to be positive and **fewer** than half to be negative:



1.3 Sign Test – test median

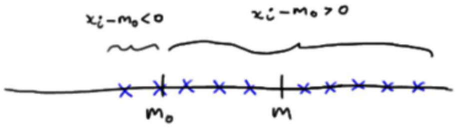
15

1. Calculate $X_i - m_0$ for $i = 1, 2, \ldots, n$.
2. Define *N*− = the number of negative signs obtained upon calculating $X_i - m_0$ for $i = 1, 2, \ldots, n$.
3. Define *N*+ = the number of positive signs obtained upon calculating $X_i - m_0$ for $i = 1, 2, \ldots, n$.

If the null hypothesis is true, then *N*− and *N*+ both follow a binomial distribution with parameters *n* and *p* = 1/2.

$$N- \sim b\left(n, \tfrac{1}{2}\right) \text{ and } N+ \sim b\left(n, \tfrac{1}{2}\right)$$

Suppose Ha: $m > m_0$,



we should reject the null hypothesis if n− is too small.

Or alternatively, if the P-value as defined by below is too small.

$$P = P(N- \leq n-)$$

联系二项分布

example

看病

1.3 Sign test for testing median: example

16

Recent studies suggested that the median length of each patient visit was 4.2 minutes. It is believed that the median visit length in practices is shorter than 4.2 minutes. A random sample of 20 visits in practices yielded the following visit lengths (minutes):

1.94 5.34 5.56 3.62 3.64 3.68 3.81 3.87 3.89 4.91
3.93 4.01 4.04 5.16 4.19 4.34 4.35 4.48 4.49 4.68

Based on these data, is there sufficient evidence to conclude that the median visit length in practices is shorter than 4.2 minutes?

清华大学统计学研究中心

注意单边检验/双边检验

sign test-summary

H0下Si的对称性非常重要

只要有等号就能放进原假设，和θ=0.5的检验和结论是同一回事？？？？？？？

Similar arguments can be used to test hypothesis

- H_0: $\theta \leq 0.5 \leftrightarrow H_1$: $\theta > 0.5$

缺点：只考虑符号，完全忽略了数值

- Drawback of sign test: it ignores magnitudes completely → it is inefficient (low power)

wilcoxon signed rank sum test

wilcoxon符号秩和检验

rank

- For example, raw data were 3.2, 2.4, 5, 3.8, the ranks were 2, 1, 4, 3
- In case of ties, mid-ranks are used, e.g., if the raw data were 105, 120, 120, 121, the ranks would be 1, 2.5, 2.5, 4

并列：midrank 均分

SR

1.5 Wilcoxon Signed Rank Sum Test

20

Subject	Drug 1	Drug 2	Diff (2-1)	Sign	Rank
1	1.9	0.7	−1.2	-	3
2	−1.6	0.8	2.4	+	8
3	−0.2	1.1	1.3	+	4.5
4	−1.2	0.1	1.3	+	4.5
5	−0.1	−0.1	0.0	NA	NA
6	3.4	4.4	1.0	+	2
7	3.7	5.5	1.8	+	7
8	0.8	1.6	0.8	+	1
9	0.0	4.6	4.6	+	9
10	2.0	3.4	1.4	+	6

Table: Hours of extra sleep on drugs 1 and 2, differences, signs and ranks of sleep study data

In the drug analysis

- Obtain S_i, the sign of $Z_i = Y_i - X_i$
- Discarding those in which $Z_i = 0$ ($S_i = 0$)
- Observations with zero differences are ignored
- Rank (R_i) = rank of $|Z_i|$ (absolute value of Z_i) after discarding $Z_i = 0$
- Signed rank: $SR = S * Rank$
- Calculate the test statistic W^+ (or W)

$$\boldsymbol{W^+ = \sum_{i=1}^{n_0} R_i I_{S_i>0}} \text{ (or } \boldsymbol{W = \sum_{i=1}^{n_0} S_i R_i}\text{)}$$

符号通过sign表示

rank：对Zi的绝对值排序

Signed rank SR=Sign * Rank

test statistic **W^+** 遍历←正数的秩 **W**

reject region

如果没有差异W+不一定很小也不一定很大（数据有限多）； W和0的差距不会太大

1.5 Wilcoxon Signed Rank Sum Test

- Under H_0 (no difference), W^+ could not be too small or too large
- Rejection H_0 if W^+ is too small or too large

```
> x <- c(1.9, -1.6, -0.2, -1.2, -0.1, 3.4, 3.7, 0.8, 0.0, 2.0)
> y <- c(0.7, 0.8, 1.1, 0.1, -0.1, 4.4, 5.5, 1.6, 4.6, 3.4)
> wilcox.test(y - x, correct = FALSE, exact = FALSE)

        Wilcoxon signed rank test

data:  y - x
V = 42, p-value = 0.02077
alternative hypothesis: true location is not equal to 0

> wilcox.test(y, x, correct = FALSE, exact = FALSE, paired = TRUE)
```

可以直接把y-x写出来做单样本test
也可以写y, x作为paired sample `paired=TRUE`