经典非参数统计

学业辅导中心

目录

1 基于二项分布的检验

1.1 二项检验*

在概率论中, 我们学过n重Bernoulli试验, 二项分布描述了n重Bernoulli试验中, 恰有k次成功的概率.

注记. 在基于二项分布的检验中, R语言代码一般看的就是`pbinom()`, `binom.test()`等.

> **例 1.** 某及其生产一种产品. 当次品率小于等于5%认为该机器工作正常. 某天抽出15件产品, 有3件次品. 问该天机器是否正常?

解答. 假设每个产品为次品的概率为p, 且是否为次品相互独立. 因此假设检验问题为

$$p \leq 0.05 \quad v.s. \quad p > 0.05$$

取检验统计量为次品的个数, 可知在零假设下, $T \sim Bernoulli(15, 0.05)$, 当$T$较大时应当拒绝$H_0$,

```
binom.test(x = 3,n = 15,p = 0.05,alternative = "greater")

    Exact binomial test

data:  3 and 15
number of successes = 3, number of trials = 15, p-value = 0.0362
alternative hypothesis: true probability of success is greater than 0.05
95 percent confidence interval:
 0.05684687 1.00000000
sample estimates:
probability of success
                   0.2
```

p值为0.0362.

1.2 符号检验

符号检验是指: 利用正负号(sign)的数目对某种假设做统计推断. 在许多情况下, 既可以使用符号检验, 也可以使用更有效的检验, 但是符号检验更简单.

备选假设	p 值	使检验有意义的条件*
$H_1 : Q_\pi > q_0$	$P_{H_0}(K \leq s^-)$	$\hat{Q}_\pi > q_0$即负号较少
$H_1 : Q_\pi < q_0$	$P_{H_0}(K \geq s^-)$	$\hat{Q}_\pi < q_0$即负号较多
$H_1 : Q_\pi \neq q_0$	$2\min\left\{P_{H_0}(K \leq s^-), P_{H_0}(K \geq s^-)\right\}$	

1.3 基于符号检验的中位数置信区间

$(X_{(k+1)}, X_{(n-k)})$, 其中k满足恰使

$$2P\left(binom\left(n, \frac{1}{2}\right) \leq k\right) \geq 0.95$$

成立的值.(95%置信区间).

1.4 小样本离散数据推断的保守性*

对于离散概率分布使用普通p值的小样本推断是保守的, 也就是说当H_0成立的时候, $p \leq 0.05$不是精确的等于0.05, 通常是小于5%. 因此第一类错误的真实概率将小于0.05.

1.5 Cox-Stuart趋势检验

1. 写出H_0, H_1
2. 算$c = \lceil \frac{n}{2} \rceil, n' = \lfloor \frac{n}{2} \rfloor$(对子数).
3. $s^+ = \sharp\{D_i : D_i = x_i - x_{i+c} > 0\}, s^- = \sharp\{D_i : D_i = x_i - x_{i+c} < 0\}$.
4. 用符号检验.

- s^+较大, 有下降的趋势; s^+较小, 有增长的趋势;
- s^-较大, 有增长的趋势; s^-较小, 有下降的趋势.

1.6 游程检验

$$P(R = 2k) = \frac{2\binom{m-1}{k-1}\binom{n-1}{k-1}}{\binom{N}{n}};$$

$$P(R = 2k+1) = \frac{\binom{m-1}{k-1}\binom{n-1}{k} + \binom{m-1}{k}\binom{n-1}{k-1}}{\binom{N}{n}}$$

- R较大: 频繁交替
- R较小: 有聚类倾向

2 基于超几何分布的检验

2.1 Brown-Mood中位数检验

1. 画出列联表:

	X	Y	总和
$> M_{XY}$	a	b	$t \equiv a+b$
$< M_{XY}$	$m-a$	$n-b$	$(m+n)-(a+b)$
总和	m	n	$N \equiv m+n$

2. $P(A=k) = \frac{\binom{m}{k}\binom{n}{t-k}}{\binom{m+n}{t}}$.

3. 计算p值: $P(A \le a)$= `phyper(a,m,n,a+b)`.

- A较大时, 说明$M_X > M_Y$更有可能;
- A较小时, 说明$M_X < M_Y$更有可能;
- 只看A的原因是给定边际后, 列联表中的数据知一求三.

2.2 Fisher精确检验

图 1: young R.A. Fisher

假定边际频数是固定的(源于女士品茶Lady Tasting Tea).

注记. *Fisher为了解释他在1935年的著作试验设计描述: 一位工作于London附近Rothamsted研究所Fisher的朋友宣称, 在饮茶时, 她能区分是先加了茶还是先加了奶. 为了检验这位朋友是否真的有这个能力, Fisher设计类一个试验, Fisher让她品尝8杯茶, 4杯先茶后奶, 4杯先奶后茶, 她被告知各有4种, 让她品尝区别.*

先加入	猜测先加入		
	奶	茶	合计
奶	*3*	*1*	*4*
茶	*1*	*1*	*4*
合计	*4*	*4*	*4*

n_{11}原假设分布服从超几何分布, 在本例中, 她猜中了四杯先加奶中的三杯, 在原假设下的分布是:

$$P(3) = \frac{\binom{4}{3}\binom{4}{1}}{\binom{8}{4}} = \frac{\frac{4!}{3!1!}}{\frac{8!}{4!4!}} = \frac{16}{70} = 0.229$$

若备择假设是猜测与顺序有正相关关系, 则一种更极端的情况是: 全猜对了. 概率为

$$P(4) = \binom{4}{4}\binom{4}{0} / \binom{8}{4} = 1/70 = 0.014.$$

同边缘分布表的超几何分布:

n_{11}	概率	P-值	X^2
0	0.014	1.000	8.0
1	0.229	0.986	2.0
2	0.514	0.757	0.0
3	0.229	0.243	2.0
4	0.014	0.014	8.0

注: P-值为单边备择假设的超几何右尾概率.

在零假设是独立/没有关系的情况下, 列联表的条件概率服从超几何分布.

$$P(n_{ij}) = \frac{\binom{n_{1\cdot}}{n_{11}}\binom{n_{2\cdot}}{n_{21}}}{\binom{n_{\cdot\cdot}}{n_{\cdot 1}}} = \frac{\binom{n_{\cdot 1}}{n_{11}}\binom{n_{\cdot 2}}{n_{12}}}{\binom{n_{\cdot\cdot}}{n_{\cdot\cdot}}} \frac{n_{\cdot 1}!n_{1\cdot}!n_{\cdot 2}!n_{2\cdot}!}{n_{\cdot\cdot}!n_{11}!n_{12}!n_{21}!n_{22}!}.$$

注记 (P值的保守性(偏小))**.** 对于小样本, 精确分布n_{11}所能取值较少, 离散性影响了假设检验的错误率, 和其它方法类似, 若设置当p值下雨等于0.05时拒绝H_0, 真实的第一类错误发生的概率可能远小于0.05. 此检验被认为是保守的, 因为实际错误率比预期小.

- 产生的若干种可能都列联表**不是等概率的**.
- 当n_{11}过小或过大时拒绝.
- 检验二维列联表的关系(即独立性)可以用Fisher精确检验.
- 代码直接用`fisher.test()`

3 基于秩的检验

3.1 秩的基本性质*

定义 (秩)**.** 设$X_1, \cdots, X_n$是样本(不必*i.i.d.*), 其值两两不同, 记

$$R_i = \sum_{j=1}^{n} I(X_j \le X_i), i = 1, \cdots, n,$$

则称R_i为X_i在样本$X_1, \cdots, X_n$中的秩.

性质 (秩的分布)**.** *1.* 设$(r_1,\cdots,r_n)$为$(1,\cdots,n)$的任一置换*(*这样的置换共有$n!$个*)*, 则

$$P\{(R_1,\cdots,R_n)=(r_1,\cdots,r_n)\}=\frac{1}{n!}$$

2. 根据抽签定理, 秩向量的边际分布是均匀分布. 特别的, 一维边际分布有

$$P(R_i=r)=\frac{1}{n},\quad r=1,\cdots,n,\quad i=1,\cdots,n$$

二维边际分布为

$$P(R_i=r,R_j=s)=\frac{1}{n(n-1)},\quad r\neq s,\quad i\neq j$$

性质 (统计推断)**.** • $\mathbb{E}R_i=\frac{n+1}{2}, i=1,\cdots,n.$

- $Var(R_i)=\frac{n^2-1}{12}, i=1,\cdots,n.$
- $cov(R_i,R_j)=-\frac{n+1}{12}, i\neq j.$

证明. 1. 由于$\sum_{i=1}^{n}R_i=\frac{n(n+1)}{2}$, 再由于$R_i$分布都是均匀分布,

$$\mathbb{E}R_i=\sum_{i=1}^{n}rP(R_i=r)=\frac{n+1}{2}$$

或者

$$n\mathbb{E}R_i=\frac{n(n+1)}{2}\Rightarrow\mathbb{E}R_i=\frac{n+1}{2}, i=1,\cdots,n.$$

2. $Var(R_i)=\mathbb{E}(R_i^2)-(\mathbb{E}R_i)^2.$

$$\mathbb{E}(R_i^2)=\sum_{i=1}^{n}r^2P(R_i=r)=\frac{1}{6}(n+1)(2n+1).$$

3. 利用$\sum_{i=1}^{n}R_i=\frac{n(n+1)}{2}$, 从而

$$0=Var(\sum_{i=1}^{n}R_i)=\sum_{i=1}^{n}Var(R_i)+\sum_{i\neq j}cov(R_i,R_j)=nVar(R_i)+(n^2-n)cov(R_i,R_j).$$

□

3.2 位置参数检验

3.2.1 Wilcoxon符号秩检验

中位数的Wilcoxon检验:

1. 求$|X_i-X|$, 并排列

2. $W^+ = \sum R_i(|X_i - X|)\{X_i > X\}, W^- = \sum R_i(|X_i - X|)\{X_i < X\}$;
3. $p = P(W < \min(W^+, W^-))$;
4. `p = psignrank(min(`W^+, W^-`), n), wilcoxon.test(y-theta, alt = "")`

- $W^+ + W^- = \dfrac{n(n+1)}{2}$
- W^+较大时, 说明对称中心应大于M_0, W^+较小时, 说明对称中心应小于M_0
- W^-较大时, 说明对称中心应小于M_0, W^-较小时, 说明对称中心应大于M_0
- Walsh平均的中位数是中位数的点估计: Hodge-Lehmann估计量.

3.2.2 成对数据的Wilcoxon符号秩检验

- 每一对数据来源于可以比较的类似的对象;
- 对和对之间是独立的;
- 都是连续型变量.

1. $D_i = X_i - Y_i$
2. $H_0 : M_D = 0$
3. $W^+, W^-, w = \min(W^+, W^-)$
4. $p = psignrank(w, n)$/`wilcox.test(x,y,paired = T, alt="")`

3.2.3 WMW秩和检验

检验两个样本所代表的中位数是否一样.

假设有m个$X : X_1, \cdots, X_m$, n个$Y : Y_1, \cdots, Y_n$, $N = m + n$

- Wilcoxon秩和统计量: $W_X = \sum_{i=1}^{m} R_i(X), W_Y = \sum_{i=1}^{n} R_i(Y)$
- Mann-Whitney统计量: $W_{XY} = \sharp\{(x_i, y_i) : x_i < y_i\}$
- 关系:
$$W_Y = W_{XY} + \frac{1}{2}n(n+1), \quad W_X = W_{YX} + \frac{1}{2}m(m+1).$$
- 代码: `wilcox.test(x,y, alt="")`

证明.

$$R_i(Y) = \sum_{j=1}^{m} I(X_j < Y_i) + \tilde{R}_i \leftarrow \text{为Y在自身的秩}$$

于是

$$W_Y = \sum_{i=1}^{n} R_i(Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} I(X_j < Y_i) + \sum_{i=1}^{n} \tilde{R}_i = W_{XY} + \frac{n(n+1)}{2}$$

□

- 当W_X较小时, $W_X = W_{YX} + \frac{1}{2}m(m+1)$较小, 认为$M_X < M_Y$;
- 当W_Y较小时, $W_Y = W_{XY} + \frac{1}{2}n(n+1)$较小, 认为$M_X > M_Y$;

精确分布:

秩	X 和 Y 的6种组合					
1	Y	Y	Y	X	X	X
2	Y	X	X	Y	Y	X
3	X	Y	X	Y	X	Y
4	X	X	Y	X	Y	Y
W_{XY}	0	1	2	2	3	4
W_{YX}	4	3	2	2	1	0
W_Y	3	4	5	5	6	7
W_X	7	6	5	5	4	3
概率	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{6}$		$\frac{1}{6}$	$\frac{1}{6}$

3.2.4 MWM统计推断*

根据前面秩的性质, 我们可以在Wilcoxon-Mann-Whitney中, 直接写出秩的统计推断性质:

- $\mathbb{E}(R_i) = \frac{n+m+1}{2}$;
- $Var(R_i) = \frac{(n+m-1)(n+m+1)}{12}$;
- $cov(R_i, R_j) = -\frac{n+m+1}{12}$.

于是

性质 (MWM统计推断)**.**

- $\mathbb{E}(W_Y) = \frac{n(n+m+1)}{2}, \mathbb{E}(W_X) = \frac{m(n+m+1)}{2}$;
- $Var(W_Y) = \frac{nm(n+m+1)}{12} = Var(W_{XY})$;
- $\mathbb{E}W_{XY} = \frac{mn}{2}$.

3.2.5 逆转WMW秩和检验得到中位数置信区间

为得到中位数差的置信区间, 我们用以下的步骤:

1. 得到mn个差$x_i - y_j$
2. 排序:$D_{(1)}, \cdots, D_{(mn)}$
3. 查$1 - 2P(W_{XY} \leq W_{\frac{\alpha}{2}}) \geq 1 - \alpha$.
4. 得到$(D_{(W_{\frac{\alpha}{2}})}, D_{(W_{mn+1-\frac{\alpha}{2}})})$

3.3 尺度参数检验

3.3.1 Siegel-Tukey方差检验

- 两独立样本, 位置参数相同(若实际位置参数不等(检验出), 可以通过平移的方式使位置参数相等).
- 基本思想是:
  1. 将两样本混合排序, 并按照特定的顺序赋予一个“秩”
  2. 按照样本, 求秩和W_X, W_Y, 计算W_{XY}, W_{YX}
  3. 令$W = \min(W_{XY}, W_{YX})$, 按照WMW分布得到结论.

- $H_a : \sigma_X > \sigma_Y$: 当$W_{YX} = W_X - \dfrac{m(m+1)}{2}$较小时, 认为$X$的方差较大;
- $H_a : \sigma_X < \sigma_Y$: 当$W_{XY} = W_Y - \dfrac{n(n+1)}{2}$较小时, 认为$Y$的方差较大;

3.3.2 Mood检验*

$$M = \sum_{j=1}^{m} \left(R_{1j} - \frac{N+1}{2} \right)^2 \ (\approx m\widehat{Var}(R(X))).$$

M较大时, 认为X的方差偏大.

```
mood.test(x, y, alternative = "")
```

3.4 多个独立样本(Rank Methods for the k-Sample Location Problem)

类似One-way anova. One-way anova的线性模型是:

$$x_{ij} = \mu + \theta_i + \varepsilon_{ij}, j = 1, ..., n_i及i = 1, ..., k,$$

上述是k组独立的随机样本, 组间和组内都是独立的, 组内还是同分布的.

3.4.1 Kruskal-Wallis检验

$$H_0 : \theta_1 = \theta_2 = \cdots = \theta_k \Leftrightarrow H_a : H_0\text{的诸等式中至少有一个不成立.}$$

混合全部数据并排秩R_{ij}, χ^2近似:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} n_i (\bar{R}_i - \bar{R})^2 \to \chi^2(k-1), \quad \min(n_1, \cdots, n_k) \to \infty.$$

F近似:

$$F_{\mathrm{R}} = \left(\frac{N-k}{k-1}\right)\left(\frac{H}{N-1-H}\right) \to F(k-1, N-k)$$

```
kruskal.test(a,b)
```

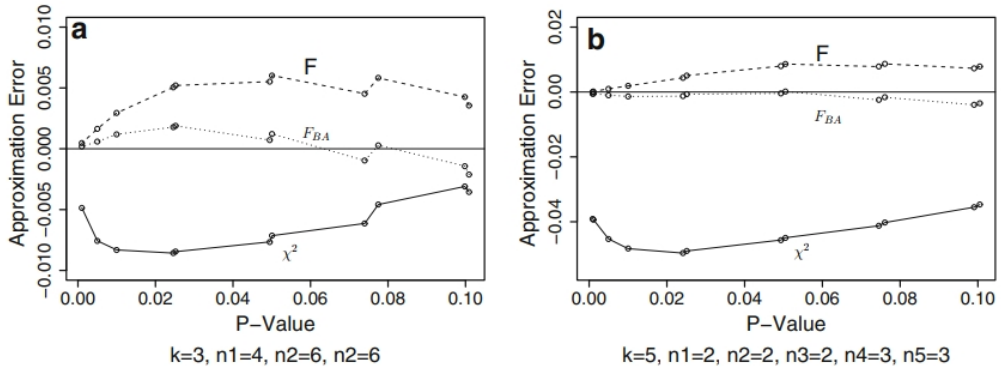


Fig. 12.4 (Exact P-Values − Approximate P-Values) versus Exact P-Values for Kruskal-Wallis Statistic. $F = F(k-1, N-k)$, $F_{BA} = F(d(k-1), d(N-k))$, and $\chi^2 = \chi^2_{k-1}$

- 小样本情况下可以求精确的p值: m/M;
- F分布比χ^2分布近似效果更好.

3.4.2 Jonckheere-Terpstra检验

相当于带趋势的单因素ANOVA: JT检验是有Jonckheere(1954)和Terpstra(1952)先后提出的, 它比KW检验有更强的功效.

$$H_0 : \theta_1 = \theta_2 = \cdots = \theta_k \Leftrightarrow H_a : \theta_1 \le \theta_2 \le \cdots \le \theta_k\text{, 且至少一个不等号是严格的.}$$

1. 写出H_0, H_1
2. 计算$U_{ij}, i < j, U_{ij} = \sum_{l=1}^{n_i} \sum_{m=1}^{n_j} I(X_{il} < X_{jm})$
3. $J = \sum_{i<j} U_{ij}$
4. 用渐进正态性验证.

3.5 区组实验设计

区组(blocking)是在试验设计中减小方差的重要方法. 常见的完全随机区组设计(Randomized Complete Block Design, RCBD)可以看成是多于两种处理的配对的情况.

完全区组设计的线性模型是:

$$Y_{ij} = \mu + \beta_i + \alpha_j + e_{ij}$$

其中α_j是处理的效应, β_i是区组的效应. 在上述记号下, 行是区组, 列是处理. 在数理统计中, 在假定误差为零均值正态分布后, 双因素ANOVA的F统计量是

$$F = \frac{\frac{1}{k-1}\sum_{j=1}^{k} n(\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot})^2}{\frac{1}{(k-1)(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{k}(Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot})^2} \sim F(k-1, (k-1)(n-1))$$

然而, F统计量在面对离群值的时候,响应变量的非正态性对第二类错误不是稳健的(比如可以尝试对Y_{ij}进行变换), 我们使用秩方法就不用对Y_{ij}的分布做假定.

3.5.1 Friedman检验(Friedman's Rank Test)

统计量:

$$Q = \frac{12}{bk(k+1)}\sum_{i=1}^{k}\left(R_i - \frac{b(k+1)}{2}\right)^2 = \frac{12}{bk(k+1)}\sum_{i=1}^{k} R_i^2 - 3b(k+1).$$

$$Q \to \chi^2(k-1)$$

- 在区组中排秩;
- 对处理求秩和.

注记. 通过输出的结果的自由度判断哪个是处理, 哪个是区组. 一般, 感兴趣的是处理.

3.5.2 Page检验

区组设计的单调备择假设, `page.trend.test()`.

构造方法:

$$L = \sum_{i=1}^{k} iR_i$$

3.5.3 Cochran检验

是一种CMH检验的退化形式,

处理	区组: 20个村民对A,B,C,D四个候选人的评价																			N_i
A	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	16
B	1	1	0	0	0	1	1	1	1	1	0	1	1	0	1	1	0	0	0	11
C	0	1	1	1	1	0	0	0	0	1	0	0	0	1	1	0	1	0	1	9
D	0	0	0	0	1	1	0	0	1	0	0	0	0	1	0	1	1	0	0	6
L_j	1	3	2	1	2	3	2	2	3	3	1	2	2	3	3	3	2	1	2	1	42

k = 4

4 相关性度量

4.1 Spearman相关检验

- 点估计:

$$r_s = \frac{\sum_{i=1}^{n}(R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^{n}(R_i - \bar{R})^2 \sum_{i=1}^{n}(S_i - \bar{S})^2}} = 1 - \frac{6\sum_{i=1}^{n}(R_i - S_i)^2}{n(n^2-1)}.$$

- 渐进正态性:$Z = r_s\sqrt{n-1} \to N(0,1)$.

4.2 Kendall τ 相关检验

从两个变量是否协同一致出发, 检验两个变量之间是否存在相关性.

- C :协同: $(x_j - x_i)(y_j - y_i) > 0$
- D :不协同:$(x_j - x_i)(y_j - y_i) < 0$
- Kendall相关系数:

$$\tau_a = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} \Psi(X_i, X_j, Y_i, Y_j) = \frac{K}{\binom{n}{2}} = \frac{n_c - n_d}{\binom{n}{2}},$$

4.3 Goodman-Kruskal's γ相关检验

有序变量的相关性检验, 与前面不同的是, 这是一个属性数据.

$$G = \frac{P - Q}{P + Q} = \frac{n_c - n_d}{n_c + n_d},$$

n_c =计数×$\sum$右下角计数, n_d =计数×$\sum$左下角计数.

5 属性数据分析

所谓列联表就是观测数据按两个或者更多属性(定性变量分类时所列出的频数表). 属性变量(categorical data)不同于连续型变量, 属性变量的观测只仅表明所属的类别. 属性数据分析方法与连续型数据分析方法差异较大, 针对属性数据分析的方法起源于20世纪英国. 20世纪早起, 属性数据主要研究变量之间的关联性.

5.1 列联表(Contingency Table)

对于单个属性变量，我们可以通过计算各个类别中观测的数目来概括数据.各个类别的样本比例估计了各类别的概率.

现假设有两个属性变量用X和Y表示. I表示的类别数而用J表示Y的类别数. 我们把 IJ种可能出现的结果放到I行J列的长方形表中, 表的I行代表X的I个水平, J列代表Y的J个水平. 表的IJ个单元代表了 IJ种可能出现的结果.

如上述形式,在每个单元里填写相应结果计数的表叫做列联表. 交叉划分两个属性变量的列联表叫做**双向列联表**; 交叉划分三个属性变量的列联表叫做**三向列联表**, 依此类推. 把一张I行J列的双向列联表称作 $I \times J$列联表.

5.1.1 联合概率, 边际概率, 条件概率

列联表的概率有两种类型:联合概率, 边际概率或条件概率. 假设目标总体按X和Y分类,

- $\pi_{ij} = P(X = i, Y = j)$为落在第$i$行, 第$j$列的概率. π_{ij}构成了X和Y的联合分布, 满足

$$\sum_{i,J} \pi_{ij} = 1$$

- 边际分布是指: 联合概率按行/列求和. 记行变量的分布为$\{\pi_{i+}\}$, 列变量的分布为$\{\pi_{+j}\}$.

- 在许多列联表中，某一个变量 (比如列变量 Y) 是响应变量，另一个变量 (行变量X) 为解释变量.那么对于X 的每个水平，我们分别构造Y 的概率分布是有意义的.这个分布由给定 X 水平下 Y 的条件概率组成，我们称其为条件分布.

5.1.2 检验诊断的灵敏度和特异度

令X表示个体疾病的真实状态, $X = 0$为阴性, $X = 1$为阳性. Y表示个体疾病的检测状态, $Y = 0$为阴性, $Y = 1$为阳性.

那么,

$$\text{敏感度 } = P(Y = 1|X = 1), \quad \text{特异度 } = P(Y = 2|X = 2).$$

灵敏度和特异度越高, 检验效果越好.

5.1.3 独立性

如果对于 X 的每个水平，Y 的条件分布是同分布的，那么我们称两个变量是统计独立的.当两个变量独立时，对于任意一列 j 各行的概率值相同.

当两个变量都是响应变量，我们可以用它们的联合分布来描述它们的关系，也可以用给定 X 时 Y 的条件分布，或者给定 Y 时 X 的条件分布来描述两个变量之间的关系.换句话说，统计独立与联合概率等于它们各自边缘概率乘积等价，

$$\pi_{ij} = \pi_{i+}\pi_{+j}, \quad \text{对于 } i = 1, \cdots, I \text{ 以及 } j = 1, \cdots, J.$$

即 X 落入第 i 行而 Y 落入第 j 列的概率等于 X 落入第 i 行的概率和 Y 落入第 j 列的概率的乘积.

5.1.4 2×2列联表中两个比例的比较

具有两个类别的属性响应变量也称作二分变量.例如，用“是”和“否”划分“是否生病”时，该属性变量就是一个二分变量. 许多研究对两个组的二分变量 Y 进行比较.数据能够列在 2×2 列联表中，其中行代表两个组而列代表响应变量 Y 的水平。本节将介绍对两个组的二分变量进行比较的度量.

注记. 书中的“按行/列结果固定”, 就是说我们感兴趣的问题是什么.

- 比例差(Difference of Proportions).

$$SE = \sqrt{\frac{\hat{\pi}_1 - (1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2 - (1-\hat{\pi}_2)}{n_2}}$$

$$(\hat{\pi}_1 - \hat{\pi}_2) \pm z_{\alpha/2}(SE)$$

例 2 (Aspirin和心脏病). 引自哈佛医学院医师健康研究课题组对阿司匹林和心肌梗死(心脏病)之间关系的报告，

```r
library(tidyverse)
asp <- tibble(Group = c("Placebo", "Placebo", "Aspirin", "Aspirin"),
                    `Myocardial Infarction` = c(" Yes", "No", " Yes", "No"),
                    Count = c(189, 10845, 104, 10933))
print(addmargins(xtabs(Count ~ Group + `Myocardial Infarction`, asp)))
p1 <- 189 / 11034
p2 <- 104 / 11037
n1 <- 11034
n2 <- 11037
se.pd <- sqrt(p1*(1-p1)/n1 + p2*(1-p2)/n2)

dif <- p1-p2
c(dif - 1.96 * se.pd, dif + 1.96 * se.pd)
```

	心肌梗死		
组	是	否	合计
安慰剂	*189*	*10845*	*11034*
阿司匹林	*104*	*10933*	*11037*

我们把表的两行当作独立二项样本. 服用安慰剂的 $n_1 = 11034$ 人中有 *189*人在研究过程中发生心肌梗死*(MI)*, 比例为 $p_1 = 189/11034 = 0.0171$. 服用阿司匹林的 $n_2 = 11037$ 人中有 *104* 人发生心肌梗死，比例为 $p_2 = 0.0094$. 样本比例差为*0.0171-0.0094=0.0077*. 比例差的标准误的估计值为

$$SE = \sqrt{\frac{(0.0171)(0.9829)}{11034} + \frac{(0.0094)(0.9906)}{11037}} = 0.0015.$$

```
> prop.test(asp$Count[c(1, 3)], asp$Count[c(2, 4)], correct = FALSE)

    2-sample test for equality of proportions without continuity
    correction

data:  asp$Count[c(1, 3)] out of asp$Count[c(2, 4)]
X-squared = 25.697, df = 1, p-value = 3.995e-07
alternative hypothesis: two.sided
95 percent confidence interval:
 0.004852875 0.010976927
sample estimates:
     prop 1      prop 2
0.017427386 0.009512485
```

比例差真值 $\pi_1 - \pi_2$ 的 *95%* 置信区间为 $0.0077 \pm 1.96(0.0015)$, 即 0.008 ± 0.003. 因为此区间只包含了正实数，我们得到 $\pi_1 - \pi_2 > 0$ 的结论，即 $\pi_1 > \pi_2$. 所以对于试验个体，服用阿司匹林减小了发生 *MI* 的风险.

- 相对风险(Ratio of Proportions/Relative Risk)

$$\text{相对风险} = \frac{\pi_1}{\pi_2}$$

例 3 (Aspirin和心脏病(cont'd))**.** 具有样本比例 p_1 和 p_2 的两个组的样本相对风险为 p_1/p_2. 对于前面的数据, 样本相对风险为 $p_1/p_2 = 0.0171/0.0094 = 1.82$. 安慰剂组发生 *MI* 的比例要高 *82%*, 样本比例差 *0.008* 使得两个组的差异仿佛微不足道. 但是相对风险却表明两组的差异在公共健康领域有着重要意义、当两个组的比例均靠近零时，仅仅通过比例差比较两个组可能会误导我们.

```
n11 <- 189
n21 <- 104
se.rr <- sqrt((1-n11/n1)/n11 + (1-n21/n2)/n21)
rr <- p1/p2
> riskscoreci(189, 11034, 104, 11037, conf.level = 0.95)

data:

95 percent confidence interval:
 1.433904 2.304713

> round(c(rr * exp(-1.96 * se.rr), rr * exp(1.96 * se.rr)), digits = 3)
[1] 1.433 2.306
```

给出的相对风险真值的 *95%*置信区间为 *(1.43,2.30)*. 我们有 *95%* 的把握相信，五年之后，服用安慰剂的病人发生 *MI* 的比例是服用阿司匹林的病人发生 *MI* 的比例的 *1.43* 到 *2.30* 倍. 这说明安慰剂组发生 *MI* 的风险至少要高 *43%*.

- 胜算比(Odds Ratio). 对于成功的概率 π, 成功的优势 (odds)定义为

$$\text{odds} = \pi/(1-\pi).$$

优势是一个非负实数，当它大于 1 时成功比失败的概率大. 当优势为 odds =4.0 时，成功的可能性是失败的 4 倍. 当成功的概率是 0.8 时，失败的概率为 0.2, 则成功的优势为 0.8/0.2=4. 于是我们预期每出现 1 次失败会有 4 次成功. 当 odds =1/4, 失败的可能性是成功的 4 倍，我们预期每出现 4 次失败会有 1 次成功.

反过来, 成功的概率是优势的函数

$$\pi = \text{odds}/(\text{odds}+1).$$

在2×2表中，第 1 行成功的优势为 $\text{odds}_1 = \pi_1/(1-\pi_1)$, 第 2 行成功的优势为$\text{odds}_2 = \pi_2/(1-\pi_2)$. 两行的优势的比值，

$$OR = \frac{\text{odds}_1}{\text{odds}_2} = \frac{\pi_1/(1-\pi_1)}{\pi_2/(1-\pi_2)},$$

称作优势比. 相对风险是两个概率的比值，而优势比 OR 是两个优势的比值.

  - 胜算比是非负实数, 胜算比为1是基准, 当胜算比位于1的两侧, 则表明了不同的关联性, 当OR>1时第1行中“成功”的优势比第 2行大.例如当OR=4时第1行中“成功”的优势是第2行“成功”的优势的4倍那么,第 1行的试验比第 2行的试验更容易成功;即$\pi_1 > \pi_2$.当OR<1时第 1行试验比第2行的试验更不容易成功, 即$\pi_1 < \pi_2$.
  - 胜算比越远离1, 代表了越强的相关性.
  - 行列颠倒后, 胜算比不变

$$OR = \frac{\pi_{11}/\pi_{12}}{\pi_{21}/\pi_{22}} = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}}.$$

  - 样本胜算比的计算:

$$\widehat{OR} = \frac{p_1/(1-p_1)}{p_2/(1-p_2)} = \frac{n_{11}/n_{12}}{n_{21}/n_{22}} = \frac{n_{11}n_{22}}{n_{12}n_{21}}.$$

样本的对数胜算比($\log(\widehat{OR})$)有更好的渐进正态性, 于是构造置信区间的时候, 先构造:

$$SE = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}.$$

$$log\hat{\theta} \pm z_{\alpha/2}(SE).$$

例 4 (Aspirin与心脏病(cont'd)). 安慰剂组的心梗的胜算估计是: $n_{11}/n_{12} = 0.0174$, 表示安慰剂组中, 每*100*个没有心梗的病人, 对应*1.74*个发生心梗的病人.

*Aspirin*组的心梗胜算的估计是: $n_{21}/n_{22} = 0.0095$, 表示处理组中, 每*100*个没有心梗的病人, 对应*0.95*个发生心梗的病人.

胜算比$\widehat{OR} = 0.0174/0.0095 = 1.832$. 即安慰剂组心梗的优势是处理组心梗优势的*1.83*倍.

$$SE = \sqrt{\frac{1}{189} + \frac{1}{10,933} + \frac{1}{104} + \frac{1}{10,845}} = 0.123$$

$$(\exp(0.365), \exp(0.846)) = (e^{0.365}, e^{0.846}) = (1.44, 2.33).$$

```
> orscoreci(189, 11034, 104, 11037, conf.level = 0.95)

data:

95 percent confidence interval:
 1.440802 2.329551
```

注记 (案例对照研究(Case-Control Studies)). 书上的中风数据, 如果按行变量固定, 结果变量就是个体以前是否中风, 这是一种回顾试的设计, 被称为*Case-control study.* 这不同于一开始给出的随机化试验*(randomized controlled trial, RCT).*

5.2 三因素列联表与Simpson's Paradox

表 2.10 死刑判决结果、被告种族和受害者种族

受害者种族	被告种族	死刑		判死刑比例
		是	否	
白人	白人	53	414	11.3
	黑人	11	37	22.9
黑人	白人	0	16	0.0
	黑人	4	139	2.8
合计	白人	53	430	11.0
	黑人	15	176	7.9

来源: M. L. Radelet and G. L. Pierce, *Florida Law Rev.*, **43**: 1-34, 1991. *Florida Law Review* 获许再版.

表 2.10 是一张 2× 2× 2 列联表—— 两行，两列，以及两层—— 它来自于种族特征对杀人被判死刑是否有影响的研究.674 个个体是 1976 年至 1987 年间佛罗里达州涉嫌杀人案件的被告. 变量 Y 是死刑判决结果，它的类别是 (是，否), X = 被告的种族，Z = 受害者的种族，它们有类别 (白人，黑人). 我们将研究被告者种族是否对死刑判决结果有影响，我们把受害者种族当作控制变量. 表 2.10 包含了受害者不同种族水平下被告者种族对死刑判决结果的两张 2×2 部分表.

5.3 $2\times 2\times k$列联表的CMH检验

Cochran-Mantel-Haenszel检验是 $2\times 2\times K$ 列联表的另一种 XY 条件独立性检验方法.该检验限定每张部分表的行总和及列总和. 于是，和费希尔精确检验一样，部分表中第 1 行第 1 列的单元计数决定了这张表其它单元的计数，在通常的抽样过程中 (例如，每张部分表的每行为二项抽样.), 限定条件下部分表 k 的第 1 行第1 列的单元计数 n_{11k} 服从超几何分布 . 检验统计量利用了每张部分表的第1行第1列单元.

$$\mu_{11k} = E(n_{11k}) = n_{1+k}n_{+1k}/n_{++k},$$

$$\mathrm{Var}(n_{11k}) = n_{1+k}n_{2+k}n_{+1k}n_{+2k}/n_{++k}^2(n_{++k}-1).$$

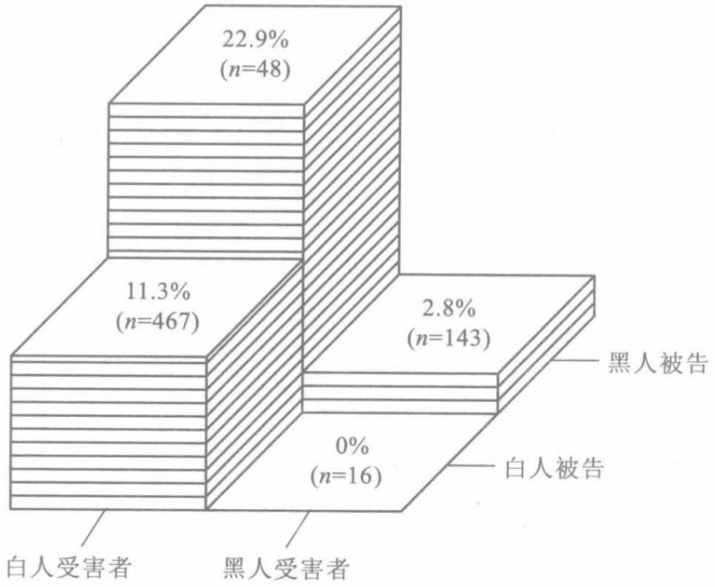


图 2.3 按被告和受害者种族划分的死刑判决比例

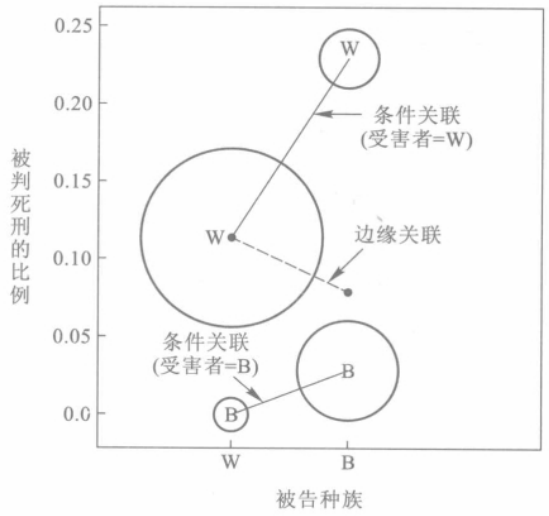


图 2.4 控制和忽略受害者种族的情况下, 不同种族的被告被判死刑的比例

检验统计量

$$\mathrm{CMH}=\frac{[\sum_k(n_{11k}-\mu_{11k})]^2}{\sum_k \mathrm{Var}(n_{11k})}$$

注记. *应用: 生存分析中的对数秩检验, 用的就是CMH检验.*

5.4 对数线性模型

为检验X,Y的独立性而建立的原假设为: 对所有的i,j, 有 $p_{ij}=p_{i\cdot}p_{\cdot j}$, 在原假设成立时, 期望频数为

$$e_{ij}=\frac{n_i\cdot n_{\cdot j}}{n}.$$

上式两边取对数可得

$$\ln e_{ij}=\ln n_{i\cdot}+\ln n_{\cdot j}-\ln n.$$

两边分别对 i,j 和 (i,j) 求和可得

$$\sum_{i=1}^{r}\ln e_{ij}=\sum_{i=1}^{r}\ln n_i+r\ln n_{\cdot j}-r\ln n,$$

$$\sum_{j=1}^{s}\ln e_{ij}=s\ln n_i+\sum_{j=1}^{s}\ln n_{\cdot j}-s\ln n,$$

$$\sum_{i=1}^{r}\sum_{j=1}^{s}\ln e_{ij}=s\sum_{i=1}^{r}\ln n_i+r\sum_{j=1}^{s}\ln n_{\cdot j}-rs\ln n.$$

可以解得

$$\ln e_{ij}=\frac{1}{s}\sum_{j=1}^{s}\ln e_{ij}+\frac{1}{r}\sum_{i=1}^{r}\ln e_{ij}-\frac{1}{rs}\sum_{i=1}^{r}\sum_{j=1}^{s}\ln e_{ij}$$

仿照方差分析模型的形式,可以将式 (4.4.5) 改写为

$$\ln e_{ij}=\mu+\lambda_i^X+\lambda_j^Y,\quad i=1,\cdots,r,j=1,\cdots,s$$

于是可以用方差分析的套路检验独立性,

模型	对数线性模型的形式	可作的检验
(X,Y,Z)	$\ln e_{ijk}=\mu+\lambda_i^X+\lambda_j^Y+\lambda_k^Z$	X,Y,Z 相互独立
(X,YZ)	$\ln e_{ijk}=\mu+\lambda_i^X+\lambda_j^Y+\lambda_k^Z+\lambda_{jk}^{YZ}$	X 与 (Y,Z) 独立
(Y,XZ)	$\ln e_{ijk}=\mu+\lambda_i^X+\lambda_j^Y+\lambda_k^Z+\lambda_{ik}^X$	Y 与 (X,Z) 独立
(Z,XY)	$\ln e_{ijk}=\mu+\lambda_i^X+\lambda_j^Y+\lambda_k^Z+\lambda_{ij}^X$	Z 与 (X,Y) 独立
(XY,XZ)	$\ln e_{ijk}=\mu+\lambda_i^X+\lambda_j^Y+\lambda_k^Z+\lambda_{ij}^{XY}+\lambda_{ik}^{XZ}$	给定 X 时 Y 与 Z 独立
(XY,YZ)	$\ln e_{ijk}=\mu+\lambda_i^X+\lambda_j^Y+\lambda_k^Z+\lambda_{ij}^XY+\lambda_{jk}^{YZ}$	给定 Y 时 X 与 Z 独立
(XZ,YZ)	$\ln e_{ijk}=\mu+\lambda_i^X+\lambda_j^Y+\lambda_k^Z+\lambda_{ik}^XZ+\lambda_{jk}^{YZ}$	给定 Z 时 X 与 Y 独立

5.5 配对数据模型

一个样本的每个个体与另一个样本的一个个体有一个自然的配对.因为一个样本中的每个观测与另一个样本中的一个观测配成对所以称两个样本的响应为配对(matched pair)由于匹配关系的存在，两个样本是统计关联的.因此.把两组观测看作独立样本的方法就不适用了.

5.5.1 McNemar检验

McNemar检验考虑的是边缘齐性的问题, 书上的例子是药的效果是否相同, 这里给一个类似的例子.

例 5 (有关环保问题的观点)**.** 对*1144*个人发放问卷, 每个人问两个问题.

付更高的税	降低生活水平		
	是	否	总和
是	*227*	*132*	*359*
否	*107*	*678*	*785*
总和	*334*	*810*	*1144*

当两个问题回答为"是"的概率相等, 此时称为是边际齐性的.

$$H_0 : \pi_{1+} = \pi_{+1}$$

由于

$$\pi_{1+} - \pi_{+1} = (\pi_{11} + \pi_{12}) - (\pi_{11} + \pi_{21}) = \pi_{12} - \pi_{21},$$

一个等价的检验是:

$$H_0 : \pi_{12} = \pi_{21}$$

渐进正态统计量是:

$$\chi = \frac{(n_{12} - n_{21})}{\sqrt{n_{12} + n_{21}}}$$

5.5.2 McNemar检验和CHM检验的关系

我们已经知道 Cochran-Mantel-Haenszel(CMH) 卡方统计量 (4.9) 检验了三向表中的条件独立性. 假设把这个统计量应用到 $2 \times 2 \times n$ 特定个体表，这个特定个体表联系了每个配对的响应结果与两个观测. 事实上，此时的 CMH 统计量代数形式上等价于 McNemar 检验统计量. 也就是说，McNemar 检验是 CMH 检验应用到 n 个配对的二分响应的特殊情况，其中 n 个配对由 n 个部分表表示.

5.5.3 分析评级者的一致性(Rater Agreement)

令 $\pi_{ij} = P(X = i, Y = j)$ 表示观察员 X 把玻片划分到类 i 且观察员 Y 把它划分到类 j 的概率.如果他们对一个特定个体划分的类别是相同的，则称他们对该个体的评级是一致的.在方形表格中，主对角线 $\{i = j\}$ 表示了观察员的一致性. 项 π_{ii} 是两个观察员均把个体划分到类 i 的概率.加和 $\sum_i \pi_{ii}$ 是一致的总概率.当 $\sum_i \pi_{ii} = 1$ 时，完全一致发生.

注记. 许多属性量表是非常主观的，所以完全一致很少出现. 这一节将介绍度量一致性的程度及探测非一致性模式的方法：一致性是与关联不同的. 强的一致性需要强的关联，但强的关联可以在没有强的一致性下存在. 若观察员 X 对个体的评级一致地比观察员 Y 高一个水平，那么一致性的强度就是较小的，尽管关联较强.

$$H_0: \text{两评委打分独立} \quad v.s. \quad H_1: \text{两评委打分一致}$$

检验统计量:

$$\kappa = \frac{p_a - p_e}{1 - p_e},$$

其中, $p_a = \sum_{i=1}^{l} n_{ii}/n, p_e = \sum_{i=1}^{l} n_{i+}n_{+i}/n^2, n = \sum_{ij} n_{ij}$.

```
> Pathology <- read.table("http://users.stat.ufl.edu/~aa/cat/data/Pathologists.dat",
+                         header = TRUE, stringsAsFactors = TRUE)
> dat <- matrix(Pathology$count, ncol = 4, byrow = TRUE)
> dat
     [,1] [,2] [,3] [,4]
[1,]   22    2    2    0
[2,]    5    7   14    0
[3,]    0    2   36    0
[4,]    0    1   17   10
> cohen.kappa(dat)
Call: cohen.kappa1(x = x, w = w, n.obs = n.obs, alpha = alpha, levels = levels)

Cohen Kappa and Weighted Kappa correlation coefficients and confidence boundaries
                 lower estimate upper
unweighted kappa  0.38     0.49  0.60
weighted kappa    0.71     0.78  0.86

 Number of subjects = 118
```

解释: 观测点一致性与独立性下所期望的一致性间的差是最大可能的差的κ.

6 CDF, ECDF与分布检验

关于累计分布函数不再赘述, 这里先回顾累积分布函数.

$$\hat{F}_n(x_0) = \frac{\text{number of } X_i \le x_0}{\text{total number of observations}} = \frac{\sum_{i=1}^{n} I(X_i \le x_0)}{n} = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le x_0)$$

这里给出一些经验累计分布函数的例子.

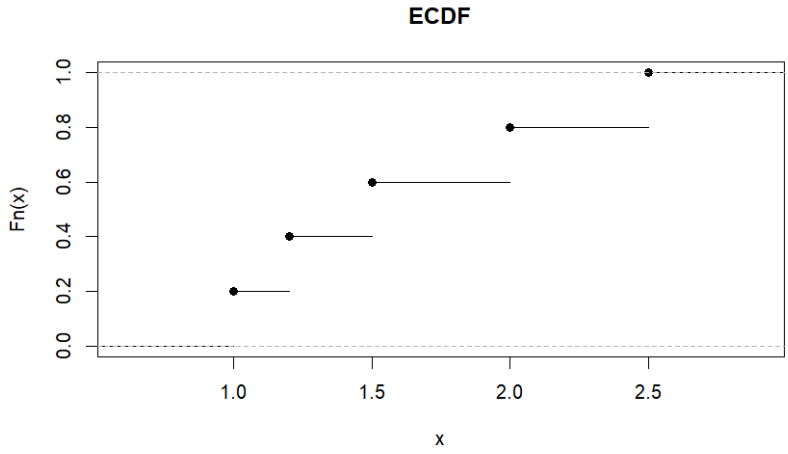
例 6. 假设观测数据是1, 1.2, 1.5, 2, 2.5, 则其*ECDF*画出图是: 代码:

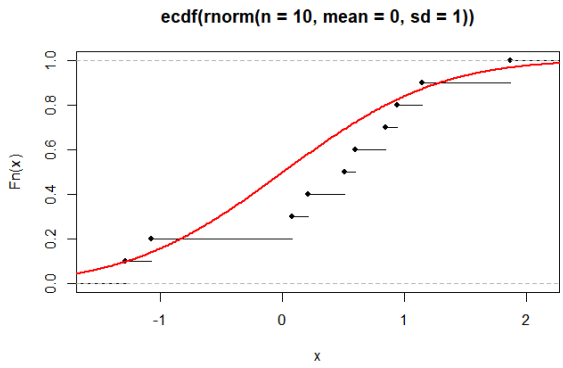
```
# Your data
data <- c(1, 1.2, 1.5, 2, 2.5)
plot(ecdf(data), main = "ECDF")
```

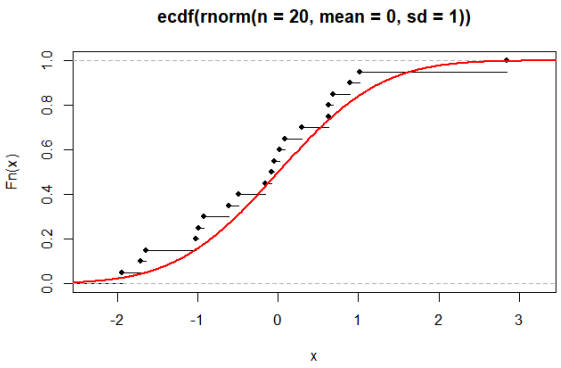
例 7. 对于随机生成的正态样本, 也可以画经验分布:

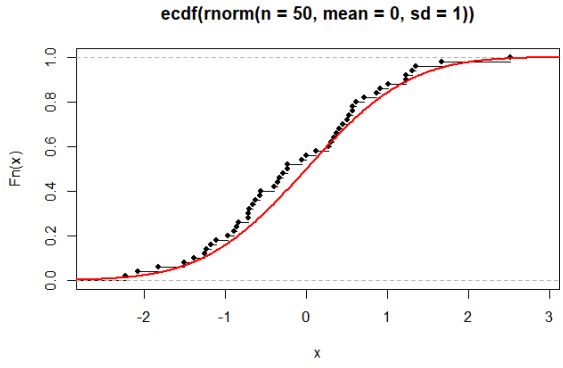
我们看到了EDF的渐进性质, 为研究EDF, 我们令

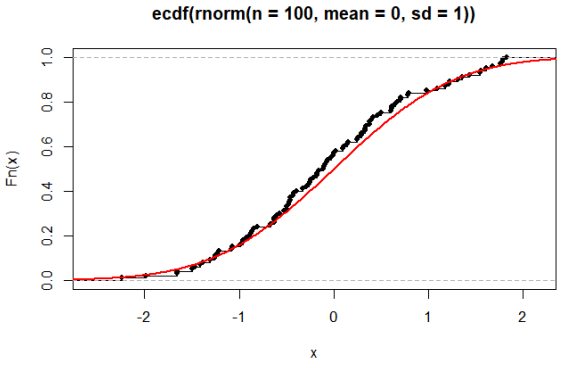
$$Y_i = \begin{cases} 1, & \text{if } X_i \le x \\ 0, & \text{if } X_i > x \end{cases}.$$


ECDF
Fn(x)
0.0
0.2
0.4
0.6
0.8
1.0
1.0
1.5
2.0
2.5
x


ecdf(rnorm(n = 10, mean = 0, sd = 1))
Fn(x)
0.0
0.2
0.4
0.6
0.8
1.0
-1
0
1
2
x


ecdf(rnorm(n = 20, mean = 0, sd = 1))
Fn(x)
0.0
0.2
0.4
0.6
0.8
1.0
-2
-1
0
1
2
3
x


ecdf(rnorm(n = 50, mean = 0, sd = 1))
Fn(x)
0.0
0.2
0.4
0.6
0.8
1.0
-2
-1
0
1
2
3
x


ecdf(rnorm(n = 100, mean = 0, sd = 1))
Fn(x)
0.0
0.2
0.4
0.6
0.8
1.0
-2
-1
0
1
2
x

从而Y_i是一个Bernoulli随机变量. 参数p满足

$$p = P(Y_i = 1) = P(X_i \le x) = F(x).$$

因此对于给定的x, 随机变量Y_i满足

$$Y_i \sim \mathsf{Ber}(F(x)).$$

这表明:

$$\mathbb{E}(I(X_i \le x)) = \mathbb{E}(Y_i) = F(x)$$
$$\mathsf{Var}(I(X_i \le x)) = \mathsf{Var}(Y_i) = F(x)(1 - F(x))$$

接下来我们考虑$\widehat{F}_n(x) = \dfrac{1}{n}\sum_{i=1}^{n} I(X_i \le x) = \dfrac{1}{n}\sum_{i=1}^{n} Y_i$. 从而

$$\mathbb{E}\left(\widehat{F}_n(x)\right) = \mathbb{E}(I(X_1 \le x)) = F(x)$$
$$\mathsf{Var}\left(\widehat{F}_n(x)\right) = \frac{\sum_{i=1}^{n}\mathsf{Var}(Y_i)}{n^2} = \frac{F(x)(1-F(x))}{n}.$$

根据LLN和CLT, 固定x, EDF是分布函数的无偏估计, 相合估计, 渐进正态估计.

$$\textbf{bias}\left(\widehat{F}_n(x)\right) = \mathbb{E}\left(\widehat{F}_n(x)\right) - F(x) = 0.$$
$$\widehat{F}_n(x) \xrightarrow{P} F(x) \Rightarrow |\widehat{F}_n(x) - F(x)| \xrightarrow{P} 0.$$
$$\sqrt{n}\left(\widehat{F}_n(x) - F(x)\right) \xrightarrow{D} N(0, F(x)(1-F(x))).$$

事实上还有更强的结论:

定理 (Glivenko-Cantelli)**.**

$$\|F_n - F\|_\infty = \sup_{x\in\mathbb{R}}\left|F_n(x) - F(x)\right| \longrightarrow 0 \quad a.s.$$

6.1 单样本Kolmogorov-Smirnov检验

我们想要假设$X_1, \cdots, X_n$是否服从某一已知分布F_0. 这是一种"拟合优度检验". 我们根据GC定理, 给出KS统计量:

$$T_{KS} = \sqrt{n}\sup_x |\widehat{F_n}(x) - F_0(x)|.$$

注记. 有多种改进形式, 见书上.

例 8. 生成的数据是50个$N(10, 25)$,检验它与$N(1, 25)$分布是否相同,

当差距足够大时, 认为两者分布不同, 如图2所示: 当红色虚线足够长就拒绝.

```
> ks.test(sample1, "pnorm", 1, 5)

    Exact one-sample Kolmogorov-Smirnov test

data:  sample1
D = 0.5856, p-value = 8.882e-16
alternative hypothesis: two-sided

```

```
sample1 <- rnorm(50, 10, 5)
sample2 <- rnorm(5000, 1, 5)
group <- c(rep("sample1", length(sample1)), rep("sample2", length(sample2)))
dat <- data.frame(KSD = c(sample1,sample2), group = group)
# create ECDF of data
cdf1 <- ecdf(sample1)
cdf2 <- ecdf(sample2)
# find min and max statistics to draw line between points of greatest distance
minMax <- seq(min(sample1, sample2), max(sample1, sample2), length.out=length(sample1))
x0 <- minMax[which( abs(cdf1(minMax) - cdf2(minMax)) == max(abs(cdf1(minMax) - cdf2(minMax))) )]
y0 <- cdf1(x0)
y1 <- cdf2(x0)
> y1- y0
[1] 0.5824
```

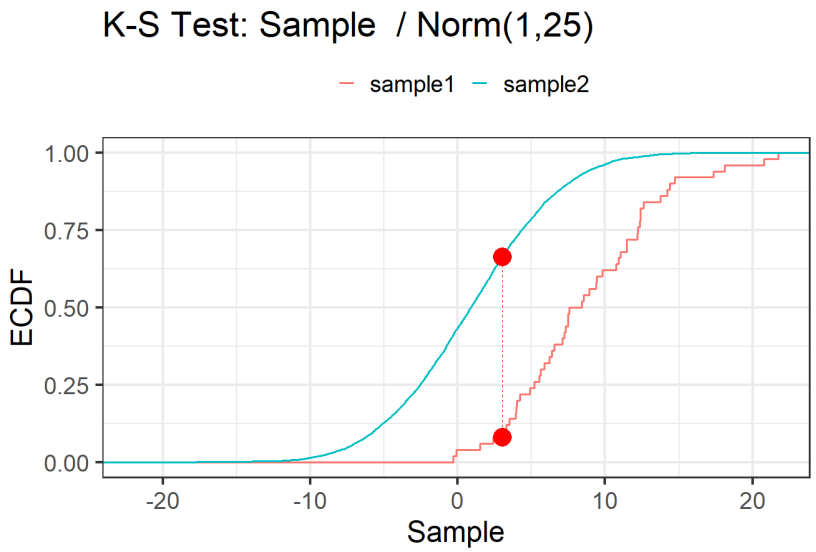


图 2: KS单样本分布检验

拒绝H_0.

注记. 还有其它的统计量: 如*Cramer-von Mises:*

$$T_{CM} = n \int \left(\widehat{F}_n(x) - F_0(x)\right)^2 dF_0(x).$$

Anderson-Darling's test :

$$T_{AD} = n \int \frac{\left(\widehat{F}_n(x) - F_0(x)\right)^2}{F_0(x)(1 - F_0(x))} dF_0(x).$$

6.2 两样本Kolmogorov-Smirnov检验

$$T_{KS} = \sqrt{\frac{nm}{n+m}} \sup_x |\widehat{F}_X(t) - \widehat{F}_Y(t)|.$$

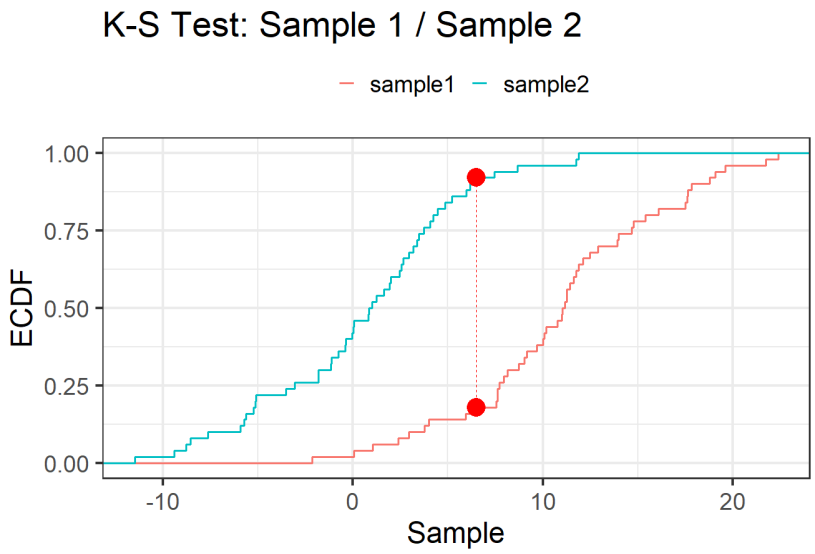


图 3: KS两样本分布检验

例 9. 类似的, 从$N(5,5)$和$N(1,5)$中各抽50个样本点, 看一下两个分布.

```
> ks.test(sample1, sample2)

    Exact two-sample Kolmogorov-Smirnov test

data:  sample1 and sample2
D = 0.58, p-value = 4.048e-08
alternative hypothesis: two-sided

> y1-y0
[1] 0.54 0.54 0.54
```

拒绝H_0.

6.3 拟合优度检验

- 二维列联表中独立性和齐性检验殊途同归;
- 多项分布+齐性检验->拟合优度检验.