- **主讲内容**
  1. LRT的渐近分布
  2. 拟合优度检验
  3. 列联表中的独立性和齐一性检验
  4. 秩检验
  5. 其它大样本检验

3.1 LRT的渐近分布

- 似然比检验统计量
$$\lambda(\mathbf{x}) = \frac{\sup_{\Theta_0} L(\theta|\mathbf{x})}{\sup_{\Theta} L(\theta|\mathbf{x})}$$
拒绝域$\{\mathbf{x} : \lambda(\mathbf{x}) \le c\}$，$0 \le c < 1$.
- 水平α检验，选择常数c，s.t.
$$\sup_{\Theta_0} \mathbb{P}_\theta(\lambda(\mathbf{X} \le c)) \le \alpha.$$
- 问题：拒绝域$\{\mathbf{x} : \lambda(\mathbf{x}) \le c\}$没有等价的简单表达，如何得出相应的样本分布进而给出水平α检验？

Theorem (定理10.3.1，【2】定理3.18，简单H_0情形)

设$X_1, \cdots, X_n$ *i.i.d.* $\sim f(x|\boldsymbol{\theta})$, $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_k)$, 参数空间$\Theta$ 是$\mathbb{R}^k$中一个含有内点的集合. 关于检验
$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \leftrightarrow H_1 : \boldsymbol{\theta} \ne \boldsymbol{\theta}_0.$$
若$f(x|\theta)$满足定理10.1.12中条件$A1 - A6$，则当H_0成立时，
$$-2\log\lambda(\mathbf{X}) \xrightarrow{L} \mathcal{X}_k^2, \quad as \quad n \to \infty.$$

复合H_0情形

- 设$X_1, \cdots, X_n$ $i.i.d. \sim f(x|\boldsymbol{\theta})$, $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_k)$, 参数空间$\Theta$ 是$\mathbb{R}^k$中一个含有内点的集合, 考虑检验问题

$$H_0 : \boldsymbol{\theta} \in \Theta_0 \leftrightarrow H_1 : \boldsymbol{\theta} \notin \Theta_0, \ \Theta_0 \subset \Theta.$$

设A是$\mathbb{R}^r$ $(r < k)$ 中一个含有内点的集合, 并在A上定义k个三阶可导函数$\mathbf{g} = (g_1, \cdots, g_k)$, 使得$A$与$\Theta_0$一一对应,

$$\Theta_0 = \{\boldsymbol{\theta} : \boldsymbol{\theta} = g(\boldsymbol{\varphi}), \boldsymbol{\varphi} \in A\}.$$

LRT统计量$\lambda(\mathbf{X}) = \frac{\prod_{i=1}^n f(X_i|\hat{\boldsymbol{\theta}}_0)}{\prod_{i=1}^n f(X_i|\hat{\boldsymbol{\theta}})}$, 这里$\hat{\boldsymbol{\theta}}_0$和$\hat{\boldsymbol{\theta}}$分别是$\boldsymbol{\theta}$在$\Theta_0$和$\Theta$上的MLE. 记$\tilde{f}(x|\boldsymbol{\varphi}) = f(x|\mathbf{g}(\boldsymbol{\varphi}))$，则可得如下结论：

Theorem (定理10.3.3，【2】定理3.19)

若$\tilde{f}(x|\boldsymbol{\varphi})$满足定理10.1.12中条件$A1 - A6$，则当$H_0$成立时，

$$-2\log\lambda(\mathbf{X}) \xrightarrow{L} \mathcal{X}^2_{k-r}, \quad as \quad n \to \infty.$$

- 【例10.3.2】 设$X_1, \cdots, X_n$ $i.i.d. \sim Poisson(\lambda)$，考虑检验问题

$$H_0 : \lambda = \lambda_0 \leftrightarrow H_1 : \lambda \neq \lambda_0,$$

请给出大样本情形下渐近水平α的LRT.

- 【例10.3.4】 设$\boldsymbol{\theta} = (p_1, p_2, p_3, p_4, p_5)$，其中$p_i > 0, i = 1, \cdots, 5$，且$\sum_{i=1}^5 p_i = 1$ 。设$X_1, \cdots, X_n$ $i.i.d.$,

$$\mathbb{P}_{\boldsymbol{\theta}}(X_i = j) = p_j, j = 1, \cdots, 5.$$

考虑检验问题

$$H_0 : p_1 = p_2 = p_3, p_4 = p_5 \leftrightarrow H_1 : H_0\text{中等式至少有一个不成立}$$

请给出大样本情形下渐近水平α的LRT。

作业 习题10：31, 33.

问题 设$X_1, \ldots, X_n$ $i.i.d. \sim F$，检验如下假设：

$$H_0 : F = F_0 \leftrightarrow H_1 : F \neq F_0 \tag{1}$$

其中F_0是一已知分布函数，称为理论分布。

- 检验的初步想法是提出一个反映实际数据与理论分布F_0偏差的统计量$D = D(X_1, \ldots, X_n; F_0)$。如果D值较大，则拒绝$H_0$.
- 如果接受H_0，可否给出刻画实际数据与理论分布F_0之间符合程度的度量？
- 令$d_0 = D(x_1, \ldots, x_n; F_0)$, 这里$(x_1, \ldots, x_n)$是样本观测值；称如下条件概率

$$p(d_0) = \mathbb{P}(D \geq d_0 | H_0)$$

为在选定的偏离指标D之下，样本与理论分布F_0的**拟合优度**(Goodness of Fit).

注1 拟合优度$p(d_0)$越接近1，表示样本与理论分布拟合得越好，因而H_0为真的可信度越高；反之，越接近0，则原假设H_0越不可信。

注2 当$p(d_0)$值低到指定水平α以下，则拒绝H_0.

- 水平为α的**拟合优度检验**

$$\text{“当}p(d_0) < \alpha\text{时，拒绝}H_0.\text{”}$$

(1) F_0离散，只取有限多个值 设$X_1, \cdots, X_n$, $i.i.d. \sim X$, $\mathbb{P}(X = a_i) = p_i$, $i = 1, \ldots, r$, i.e. 理论分布

X	a_1	a_2	$\cdots$	a_r
$\mathbb{P}$	p_1	p_2	$\cdots$	p_r

- 记$X_1, \cdots, X_n$ 中等于a_i的个数为ν_i, $i = 1, \ldots, r$, 称为观察频数，显然$\sum_{i=1}^{r} \nu_i = n$;

注 p_i的MLE$\hat{p}_i = \frac{\nu_i}{n}$，称$a_i$的频率，而称$np_i$为$a_i$的理论频数。

拟合优度检验：离散情形

- 样本与理论分布的偏离指标(K. Pearson,1900):

$$K_n = K_n(X_1, \ldots, X_n; F_0) = \sum_{i=1}^{r} \frac{(\nu_i - np_i)^2}{np_i}$$

注 记$\lambda(\mathbf{X})$是当前情形下检验问题(1)的LRT统计量. 可求得$\hat{p}_i = \frac{\nu_i}{n}$是$p_i$的MLE. 验证

$$-2\log\lambda(\mathbf{X}) = 2\sum_{i=1}^{r} \nu_i \log\frac{\hat{p}_i}{p_i} = K_n + op(1).$$

从而可得：

Theorem (3.1, K.Pearson)

(【0】定理6.4.1) 若原假设H_0成立，则当样本容量$n \to \infty$ 时，有

$$K_n \xrightarrow{L} \chi^2_{r-1}.$$

- 水平近似为α的检验：

$$\text{“当}K_n > \chi^2_{r-1}(\alpha)\text{时拒绝}H_0.\text{”}$$

改进 样本与理论分布偏离程度的度量：**拟合优度**

$$p(k_0) = \mathbb{P}(K_n \geq k_0 | H_0) \approx \mathbb{P}(\chi^2_{r-1} \geq k_0),$$

注 $p(k_0)$值越大，则接受H_0的判定可靠性越高。

Example (3.2.1)

现有属于5条生产线的不合格产品个数分别为15，27，31，19，11，问这5条生产线的不合格产品所占比率是否相等？检验水平$\alpha = 0.05$.

- 检验假设

$$H_0 : p_1 = p_2 = \cdots = p_5 = 0.2 \leftrightarrow H_1 : \text{至少有两个}p_i, p_j\text{不相等}.$$

- 水平为α的拟合优度检验:

$$p(k_0) = \mathbb{P}(K_n \geq k_0 | H_0) \approx \mathbb{P}(\chi^2_4 \geq 13.36) < 0.01 < \alpha,$$

因此拒绝H_0.

练习 【0】例6.4.1

作业 【0】习题6：Ex. 7, 8.

拟合优度检验：连续情形

(2) F连续，把c.d.f.分割成若干份，化成离散型

- 设$X_1, \cdots, X_n$, $i.i.d. \sim X$, X样本空间R_m, $m \geq 1$。

Step1 将R_m分解为r个彼此无公共点的区间或区域$I_1, \ldots, I_r$.

Step2 计算r个事件在H_0成立下的概率

$$p_j = \mathbb{P}_F(X \in I_j), j = 1, \ldots, r.$$

Step3 令ν_j为$X_1, \ldots, X_n$中落入I_j的观察频数，$j = 1, \ldots, r$. 计算

$$K_n = \sum_{i=1}^{r} \frac{(\nu_i - np_i)^2}{np_i}$$

Step4 利用定理3.1，得水平近似为α的检验

“当$K_n > \chi^2_{r-1}(\alpha)$时拒绝$H_0$.”

以及拟合优度$p(k_0)$.

3.2.2 理论分布带有未知参数情形下的拟合优度

问题 设$X_1, \cdots, X_n$ $i.i.d. \sim X$, 检验如下假设：

$$H_0: \text{存在}\boldsymbol{\theta}_0 \in \Theta_0, \text{使得X的分布为}F(x; \boldsymbol{\theta}_0). \tag{2}$$

其中$F(x; \boldsymbol{\theta})$是带有未知参数$\boldsymbol{\theta} = (\theta_1, \cdots, \theta_s) \in \Theta_0$的一个确定分布族。

(1) F离散，只取有限多个值 设$X_1, \cdots, X_n$, $i.i.d. \sim X$, $\mathbb{P}(X = a_i) = p_i(\boldsymbol{\theta})$, $i = 1, \ldots, r$, i.e. 理论分布

X	a_1	a_2	$\cdots$	a_r
$\mathbb{P}$	$p_1(\boldsymbol{\theta})$	$p_2(\boldsymbol{\theta})$	$\cdots$	$p_r(\boldsymbol{\theta})$

Step1 求$\boldsymbol{\theta}$的MLE $\widehat{\boldsymbol{\theta}}$，它是似然方程$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{0}$ 的唯一解，其中

$$\ell(\boldsymbol{\theta}) = \ln L, \quad L = \frac{n!}{\nu_1! \cdots \nu_r!} \Big(p_1(\boldsymbol{\theta})\Big)^{\nu_1} \cdots \Big(p_r(\boldsymbol{\theta})\Big)^{\nu_r}.$$

拟合优度检验：离散情形

注　这里假设$p_j(\boldsymbol{\theta}), j = 1, \ldots, r$满足如下两个条件：

1. $\sum_{j=1}^r p_j(\boldsymbol{\theta}) = 1$，$\forall \boldsymbol{\theta} \in \Theta_0$;
2. $p_j(\boldsymbol{\theta})$关于$\boldsymbol{\theta}$有一阶连续偏导，$\forall j = 1, \ldots, r$.

- 检验问题(2)可化为如下检验问题

$$H_0 : \mathbb{P}(X = a_j) = p_j(\widehat{\boldsymbol{\theta}}), \quad j = 1, \cdots, r.$$

Step2　给出检验统计量

$$K_n^* = \sum_{j=1}^r \frac{(\nu_j - np_j(\widehat{\boldsymbol{\theta}}))^2}{np_j(\widehat{\boldsymbol{\theta}})}$$

这里，ν_j是a_j的观察频数, $j = 1, \cdots, r$.

注　记$\lambda(\mathbf{X})$是当前情形下检验问题(2)的LRT统计量. 验证

$$-2 \log \lambda(\mathbf{X}) = 2 \sum_{i=1}^r \nu_i \log \frac{\hat{p}_i}{p_j(\widehat{\boldsymbol{\theta}})} = K_n^* + op(1),$$

其中$\hat{p}_i = \frac{\nu_i}{n}$. 由定理10.3.3可得如下结果：

Theorem (3.2, R.A. Fisher)

(【0】定理6.4.2) 设检验问题(2)原假设H_0成立，若$\hat{\theta}$是$\vec{\theta}$的相合估计，则当$n \to \infty$时，

$$K_n^* \xrightarrow{L} \chi^2_{r-1-s}.$$

- 检验问题(2)水平近似为α的检验：

$$当K_n^* > \chi^2_{r-1-s}(\alpha)时，拒绝H_0.$$

Step3 记k_0^*为样本算得的K_n^*具体值，检验的拟合优度

$$p(k_0^*) = \mathbb{P}(K_n^* \geq k_0^* | H_0) \approx \mathbb{P}(\chi^2_{r-1-s} \geq k_0^*).$$

- 对于给定的检验水平α，若$p(k_0^*) < \alpha$，则拒绝H_0.

(2) 当F为连续情形时，检验方法同3.2.1小节，把c.d.f.分割成若干份，化成离散型。

Example (3.2.2)

有1000人按性别与色盲分类如下

	正常	色盲
男	442	38
女	514	6

按遗传学模型，数据应有下列相对应的概率

	正常	色盲
男	$p/2$	$(1-p)/2$
女	$\frac{1}{2}p^2+p(1-p)$	$\frac{1}{2}(1-p)^2$

其中$p\in[0,1]$. 问数据与模型是否相符？检验水平$\alpha=0.05$.

举例说明

Step1 求未知参数p的MLE，得$\hat{p} \approx 0.913$。

- 问题化为检验

$$H_0 : p_{11}^* = 0.4565, \quad p_{12}^* = 0.0435, \quad p_{21}^* = 0.496, \quad p_{22}^* = 0.004,$$

Step2 求K_n^*的样本值，得$k_0^* = 2.81$.

Step3 计算拟合优度。这里r=4, s=1, 故r-1-s=2，则

$$p(k_0^*) = \mathbb{P}(K_n^* \geq k_0^* | H_0) \approx \mathbb{P}(\chi_2^2 \geq 2.81) > 0.10 > \alpha,$$

不拒绝H_0，即没有充分的证据表明数据与模型不相符。

练习 【0】例6.4.2, 6.4.3.

作业 【0】习题6，Ex.10, 12.

问题 设某总体内的每个个体（事物或人）的两个属性A和B，各有r和s个水平，检验这两种属性的关联性。

Example (3.3.1, 吸烟和肺癌的关系)

设总体为某地区特定的一群人，从总体中随机抽取n个人作调查，结果如下

A ＼ B	患肺癌	未患肺癌
吸烟	n_{11}	n_{12}
不吸烟	n_{21}	n_{22}

问吸烟是否会导致肺癌？

解答 令$X = \begin{cases} 1, & \text{吸烟} \\ 0, & \text{不吸烟} \end{cases}$，$Y = \begin{cases} 1, & \text{患肺癌} \\ 0, & \text{未患肺癌} \end{cases}$。若吸烟与肺癌无关，即$X$与$Y$之间相互独立。

- 因此，问题等价于检验

$$H_0 : \mathbb{P}(X = i, Y = j) = \mathbb{P}(X = i)\mathbb{P}(Y = j), \quad i, j = 0, 1.$$

- 设属性A,B分别有r和s个水平，分别记为$A_1, \ldots, A_r, B_1, \ldots, B_s$。
- 引进随机向量$(X^{(1)}, X^{(2)})$分别记同一个体的A,B属性的水平。第i个个体的观察结果记为$X_i = (X_i^{(1)}, X_i^{(2)})$。
- $r \times s$列联表（Contingency Table）

$X^{(1)} \diagdown X^{(2)}$	1	$\cdots$	j	$\cdots$	s	$\sum$
1	n_{11}	$\cdots$	n_{1j}	$\cdots$	n_{1s}	$n_{1\cdot}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
i	n_{i1}	$\cdots$	n_{ij}	$\cdots$	n_{is}	$n_{i\cdot}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
r	n_{r1}	$\cdots$	n_{rj}	$\cdots$	n_{rs}	$n_{r\cdot}$
$\sum$	$n_{\cdot 1}$	$\cdots$	$n_{\cdot j}$	$\cdots$	$n_{\cdot s}$	n

- $n_{i\cdot} = \sum_j n_{ij}$，$n_{\cdot j} = \sum_i n_{ij}$，$n = \sum_i n_{i\cdot} = \sum_j n_{\cdot j}$.

问题 检验

$$H_0 : X^{(1)}和X^{(2)}独立 \tag{3}$$

- 记$\mathbb{P}(X^{(1)} = i, X^{(2)} = j) = p_{ij}, \quad i = 1, \cdots, r, \quad j = 1, \cdots, s,$

$$p_{i\cdot} = \sum_{j=1}^{s} p_{ij}, \quad p_{\cdot j} = \sum_{i=1}^{r} p_{ij}.$$

- 则检验问题(3)等价于

$$H_0 : \mathbb{P}(X^{(1)} = i, X^{(2)} = j) = p_{i\cdot}p_{\cdot j}, \quad i = 1, \cdots, r, j = 1, \cdots, s \tag{4}$$

满足约束条件

$$\sum_{i=1}^{r} p_{i\cdot} = \sum_{j=1}^{s} p_{\cdot j} = 1.$$

- 用极大似然估计法求得$p_{i\cdot}$和$p_{\cdot j}$的MLE如下：

$$\begin{aligned} \hat{p}_{i\cdot}^{*} &= \frac{n_{i\cdot}}{n}, \quad i = 1, \cdots, r, \\ \hat{p}_{\cdot j}^{*} &= \frac{n_{\cdot j}}{n}, \quad j = 1, \cdots, s. \end{aligned}$$

- 检验问题(4)可化为如下检验问题　（参考【0】6.4.3）

$$H_0 : \mathbb{P}(X^{(1)} = i, X^{(2)} = j) = \hat{p}^*_{i\cdot}\hat{p}^*_{\cdot j}, \quad i = 1, \cdots, r, j = 1, \cdots, s$$

- 检验统计量

$$\begin{aligned} K_n^* &= \sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(n_{ij} - n\hat{p}^*_{i\cdot}\hat{p}^*_{\cdot j})^2}{n\hat{p}^*_{i\cdot}\hat{p}^*_{\cdot j}} = \sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(n_{ij} - n_{i\cdot}n_{\cdot j}/n)^2}{n_{i\cdot}n_{\cdot j}/n} \\ &= n\left(\sum_{i=1}^{r}\sum_{j=1}^{s}\frac{n_{ij}^2}{n_{i\cdot}n_{\cdot j}} - 1\right) \end{aligned}$$

- 类似【0】定理6.4.2，可证得当H_0成立时，有

$$K_n^* \xrightarrow{L} \chi^2_{rs-1-(r+s-2)} = \chi^2_{(r-1)(s-1)}, \quad \text{as} \quad n \to \infty.$$

- 检验问题(3)水平近似为α的检验：

$$当 K_n^* > \chi^2_{(r-1)(s-1)}(\alpha)时，拒绝 H_0.$$

- 记k_0^*为样本算得的K_n^*具体值，检验的拟合优度

$$p(k_0^*) = \mathbb{P}(K_n^* \geq k_0^* | H_0) \approx \mathbb{P}(\chi^2_{(r-1)(s-1)} \geq k_0^*).$$

- 拟合优度检验：若$p(k_0^*) < \alpha$，则拒绝H_0.

Example (3.3.2)

下表是一份寄往150个家庭的问卷调查汇总表，这份调查是某地发出要建一个污染比较严重的工厂的通告后，咨询居民对于疏散的态度。将居民根据社区与工厂距离以及他们对于疏散的意见分类如下：

疏散＼距离	1-6	7-12	13+	合计
同意	18	15	33	66
反对	20	19	45	84
合计	38	34	78	150

问：居民对于疏散的态度是否依赖于居住地离工厂的距离？请给出拟合优度检验结果，检验水平$\alpha = 0.05$。

练习 【0】例6.5.1.

作业 【0】习题6，Ex.16

3.3.2 列联表中的齐一性检验

问题 r个工厂生产同一种产品，分等级$1, \cdots, s$。记$p_i(j)$为第i个工厂生产j等产品比率，$i = 1, \cdots, r$，$j = 1, \ldots, s$。问这r个工厂产品同一等级质量是否相同？

解释 问题即为检验“r个工厂产品质量齐一性”这一假设，可表示如下：

$$H_0 : p_1(j) = p_2(j) = \cdots = p_r(j), \quad j = 1, \cdots, s. \tag{5}$$

解决 设从第i个工厂随机抽取$n_{i\cdot}$个产品，$1, \cdots, s$等品个数分别为$n_{i1}, \cdots, n_{is}$。标记工厂为X，等级为a，则可得表如下：

$X \diagdown a$	1	$\cdots$	j	$\cdots$	s	$\sum$
1	n_{11}	$\cdots$	n_{1j}	$\cdots$	n_{1s}	$n_{1\cdot}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
i	n_{i1}	$\cdots$	n_{ij}	$\cdots$	n_{is}	$n_{i\cdot}$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
r	n_{r1}	$\cdots$	n_{rj}	$\cdots$	n_{rs}	$n_{r\cdot}$
$\sum$	$n_{\cdot 1}$	$\cdots$	$n_{\cdot j}$	$\cdots$	$n_{\cdot s}$	n

注 此表与本章4.1小节中独立性检验的列联表相同，区别在于4.1小节的列联表中$n_{i\cdot}$是随机的，依赖于抽样，而此表的$n_{i\cdot}$是事先选定的已知数。

Case I 分布已知

$$p_1(j) = p_2(j) = \cdots = p_r(j) = p_j^0, \quad j = 1, \cdots, s.$$

这里$p_1^0, \cdots, p_s^0$均已知，满足$\sum_{j=1}^{s} = 1$。

- 检验统计量

$$K_n = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{(n_{ij} - n_{i\cdot} p_j^0)^2}{n_{i\cdot} p_j^0}.$$

- 类似【0】定理6.4.1，可证得当H_0成立时，有

$$K_n \xrightarrow{L} \chi^2_{rs-r}, \quad as \quad n_{i\cdot} \to \infty \quad for \quad every \quad i = 1, \ldots, r.$$

- 检验问题(5)水平近似为α的检验：

当$K_n > \chi^2_{(s-1)r}(\alpha)$时，拒绝$H_0$.

- 记k_0为样本算得的K_n具体值，检验的拟合优度

$$p(k_0) = \mathbb{P}(K_n \geq k_0 | H_0) \approx \mathbb{P}(\chi^2_{(s-1)r} \geq k_0).$$

若$p(k_0^*) < \alpha$，则拒绝H_0.

Case II 分布未知

$$p_1(j) = p_2(j) = \cdots = p_r(j) = p_j, \quad j = 1, \cdots, s.$$

这里$p_1, \cdots, p_s$皆未知。

- $p_1, \cdots, p_s$的MLE:

$$\hat{p}_j^* = \frac{n_{\cdot j}}{n}, \quad j = 1, \cdots, s.$$

- 代入K_n，得检验统计量

$$K_n^* = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{(n_{ij} - n_{i\cdot} n_{\cdot j}/n)^2}{n_{i\cdot} n_{\cdot j}/n} = n \left(\sum_{i=1}^{r} \sum_{j=1}^{s} \frac{n_{ij}^2}{n_{i\cdot} n_{\cdot j}} - 1 \right)$$

- 当H_0成立时，有 (参考【0】p263)

$$K_n^* \xrightarrow{L} \chi^2_{(r-1)(s-1)}, \quad \text{as} \quad n_{i\cdot} \to \infty \quad \textit{for} \quad \textit{every} \quad i = 1, \ldots, r.$$

举例说明

- 检验问题(5)水平近似为α的检验：
$$\text{当}K_n^* > \chi^2_{(r-1)(s-1)}(\alpha)\text{时，拒绝}H_0.$$
- 记k_0^*为样本算得的K_n^*具体值，检验的拟合优度
$$p(k_0^*) = \mathbb{P}(K_n^* \geq k_0^* | H_0) \approx \mathbb{P}(\chi^2_{(r-1)(s-1)} \geq k_0^*).$$
若$p(k_0^*) < \alpha$，则拒绝H_0.

Example (3.3.3)

为了比较三条生产线生产的不合格产品比率，质量控制工程师从每条生产线随机抽取500个产品，三条生产线的不合格数量分别为12，17，7。问三条生产线的产品合格率有差异吗？生产线和不合格情况相关吗？请给出拟合优度检验结果，检验水平$\alpha = 0.05$。

举例说明

提示 检验假设$H_0 : p_1 = p_2 = p_3$.

- 计算K_n^*的样本值得$k_0^* = 4.269$，从而拟合优度

$$p(k_0^*) = \mathbb{P}(K_n^* \geq k_0^* | H_0) \approx \mathbb{P}(\chi_2^2 \geq 4.269) > 0.10 > \alpha.$$

因此，没有充分证据表明三条生产线的不合格率存在差异。

练习 【0】例6.5.2.

作业 【0】习题6，Ex.17.