第八章　假设检验

- 为了推断总体的某些未知特性,提出关于总体的某种推断或猜测，即假设；然后通过试验，抽取样本，根据样本信息对“假设”的正确性进行判断，即检验。
- 假设检验包括：对总体参数的假设检验；
  对总体分布的假设检验；
  独立性检验等等。

§ 8.1 假设检验的基本概念

- 一. 问题的提出
- 例8.1.1：某厂生产的一种保健食品.已知在正常的情况下,每瓶保健品的重量(单位:千克)服从均值为25.0的正态分布（标准差为0.1）.某天开工后,随机抽取9瓶,测得其平均重量为24.94,试问该天生产是否正常?

$$H_0: \mu = 25; \qquad H_1: \mu \neq 25$$

- 例8.1.2：某厂生产一批产品，要求次品率不超过5%。随机抽取50件，发现有4件次品，问产品能否出厂？

$H_0: p \le 0.05$;　　$H_1: p > 0.05$

- 例8.1.3：抽取某校高数试卷60份，得60个数据。问能否由此认为成绩服从正态分布？

H_0: X服从正态分布

- 假设：根据实际问题的要求而提出的关于总体的某个命题，称为假设，记为H。
- 假设分为原假设（或零假设，记为H_0）和备择假设（或对立假设，记为H_1）。

- 检验：根据样本观测值来判断H_0的真伪的过程称为检验。若作出“H_0不成立”的结论，则称拒绝H_0（或接受H_1）；反之则称不拒绝H_0（或接受H_0或拒绝H_1）。

- 二. 假设检验基本步骤
- (1) 根据实际问题提出原假设和备择假设;
- (2) 根据原假设构造一个合适的统计量，满足在原假设为真时，统计量的分布已知且与参数无关;
- (3) 根据给定的显著性水平，确定临界值和拒绝域（对原假设不利的小概率事件）;
- (4) 根据样本观察值，计算统计量的值，并做出判断。

- 例8.1.1：某厂生产的一种保健食品.已知在正常的情况下,每瓶保健品的重量(单位:千克)服从均值为25.0的正态分布（ 标准差为0.1 ）.某天开工后,随机抽取9瓶,测得其平均重量为24.94,试问该天生产是否正常?
- 1. 提出假设

$$H_0:\mu=25; \qquad H_1:\mu\neq 25$$

- 2. 构造统计量

若原假设为真，样本均值与25的差距不会太大。选取合适的k，当$|\bar{X}-25|>k$时，拒绝原假设H_0

- 令 $$U=\frac{\bar{X}-25}{\sigma/\sqrt{n}}$$

其中σ=0.1，当H_0为真时, U~$N(0,1)$。

- 3. 给定一个小概率α(显著性水平，即犯错的概率),
$$P(|\bar{X}-25|>k)=P(|U|>k\frac{\sqrt{n}}{\sigma})=\alpha$$

- 即 $$k=u_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

临界值

- 于是得到拒绝域:
$$W=\{U:|U|>u_{\frac{\alpha}{2}}\}=\left\{(X_1,\cdots,X_n):\left|\frac{\bar{X}-25}{\sigma/\sqrt{n}}\right|>u_{\frac{\alpha}{2}}\right\}$$

- 接受域：

$$\{|U| \le u_{\frac{\alpha}{2}}\}$$

- 若取α=0.05，则拒绝域：

$$W = \left\{ |U| = \left| \frac{\bar{X} - 25}{\sigma / \sqrt{n}} \right| > u_{\frac{\alpha}{2}} = 1.96 \right\}$$

- 4.而

$$\bar{x} = 24.94, |u| = \frac{|24.94 - 25|}{0.1 / \sqrt{9}} = 1.8 < 1.96$$

所以接受原假设，即认为该日生产正常。

- 三.假设检验中的两类错误

客观 / 主观	H_0真	H_0不真
拒绝H_0	第一类错误(弃真)	
不拒绝H_0		第二类错误(存伪)

- 一般情况下，犯两类错误的概率存在此消彼长的关系，不能同时达到最小。**奈曼-皮尔逊原则是首先控制犯第一类错误的概率≤α，然后尽量降低犯第二类错误的概率。**

- 第一类错误与第二类错误的计算

$$P(\text{拒绝}H_0 \mid H_0\text{真}) = \mathrm{P}(\text{第一类错误})=\alpha$$

$$P(\text{不拒绝}H_0 \mid H_0\text{不真}) = \mathrm{P}(\text{第二类错误})=\beta$$

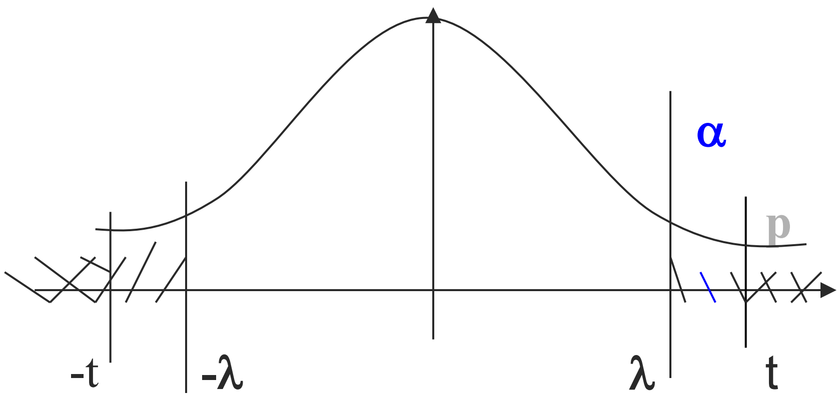
- 第一类错误即显著性水平；第二类错误较为复杂。以例8.1.1为例：

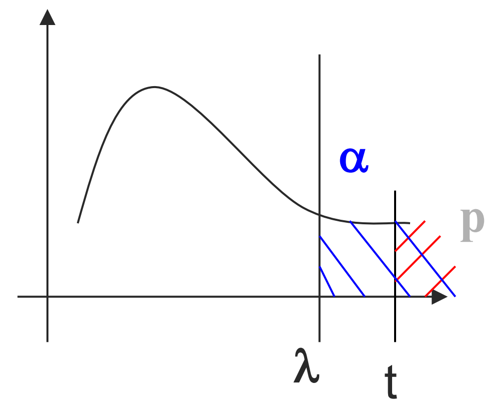
$$\beta(\mu_1) = P(\text{不拒绝}H_0 \mid H_0\text{不真}) = P(|U| < u_{\alpha/2} \mid \mu = \mu_1)$$

$$= \Phi(u_{\alpha/2} - \sqrt{n}(\mu_1 - 25)/0.1) - \Phi(-u_{\alpha/2} - \sqrt{n}(\mu_1 - 25)/0.1)$$

四. p-值检验法

p-值的定义：假设在检验时所考虑的拒绝域为W={|T|> λ}，若根据样本观测值计算出来的统计量值为t，则p=P(|T|>t)。当给定的显著性水平大于p-值时，拒绝原假设，否则不拒绝原假设。即**p-值越小越拒绝原假设**。





假设检验的三种方法：

1. 构造拒绝域；
2. 寻找置信区间；
3. 计算p-值（多用于计算机）

§ 8.2 正态总体均值的检验

以下均假设显著性水平为α。

一. 单个正态总体均值的检验

设总体$X \sim N(\mu, \sigma^2)$, $X_1, X_2, \ldots, X_n$是来自X的样本,

1.方差σ^2已知, 检验$H_0: \mu = \mu_0$; $H_1: \mu \neq \mu_0$ —u检验法

构造统计量: $$U = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

当H_0为真时, $U \sim N(0,1)$

|U|越大，对原假设越不利，故拒绝域应是形为{|U|>c}的事件。

令 $$W = \left\{ |U| = \left| \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \right| > u_{\frac{\alpha}{2}} \right\}$$

则 $P_{H_0}(W) = \alpha$ 故W为水平α的拒绝域。

考虑总体均值μ的点估计，则主元为

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$$

由 $$\left| \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \right| \le u_{\frac{\alpha}{2}}$$

可得区间估计。

方差已知时的单边假设检验

(1)方差σ^2已知，检验H_0: $\mu=\mu_0$; H_1: $\mu>\mu_0$

检验统计量同前： $U=\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$

当H_0为真时，$U\sim N(0,1)$

U越大，对原假设越不利，故拒绝域应是形为$\{U>c\}$的小概率事件。

又因为, $P_{H_0}\left(\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\geq u_\alpha\right)=\alpha$

故 $W=\{U>u_\alpha\}$ 为水平α的拒绝域。

(2)方差σ^2已知，检验$H_0:\mu=\mu_0$; $H_1: \mu<\mu_0$

检验统计量同前： $U=\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$

U越小，对原假设越不利，故拒绝域应是形为$\{U<c\}$的小概率事件。

注意到 $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathrm{N}(0,1)$

因此，$P_{H_0}\left(\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}<-u_\alpha\right)=\alpha$

对应的拒绝域为： $W=\left\{U<-u_\alpha\right\}$

2.方差σ^2未知，检验$H_0:\mu=\mu_0$; $H_1: \mu\neq\mu_0$ —t 检验法

构造统计量 $$T=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}$$

当H_0为真时, $T\sim t(n\text{-}1)$

|T|越大,对原假设越不利,故拒绝域应是形为{|T|>c}的小概率事件。

对给定显著性水平α，检验拒绝域为

$$W=\left\{|T|=\left|\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\right|>t_{\frac{\alpha}{2}}(n-1)\right\}$$

考虑总体均值 μ 的区间估计，则主元为

$$T=\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$$

由$\left|\frac{\bar{X}-\mu}{S/\sqrt{n}}\right|\le t_{\frac{\alpha}{2}}(n-1)$

可得 μ 的区间估计

方差未知时的单边假设检验

(1)检验H_0:$\mu \le \mu_0$(或$\mu = \mu_0$)；　H_1: $\mu > \mu_0$

统计量　$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

拒绝域:　$$W = \left\{ \mathrm{T} > t_\alpha(n-1) \right\}$$

(2)检验H_0:$\mu \ge \mu_0$(或$\mu = \mu_0$)；　H_1: $\mu < \mu_0$

统计量　$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

拒绝域:　$$W = \left\{ \mathrm{T} < -t_\alpha(n-1) \right\}$$

二. 两个正态总体均值的检验

$X \sim N(\mu_1, \sigma_1^2), X_1, \cdots, X_{n_1}$ $\quad Y \sim N(\mu_2, \sigma_2^2), Y_1, \cdots, Y_{n_2}$，两样本相互独立

1.当σ_1^2, σ_2^2已知时，检验H_0: $\mu_1 = \mu_2$; H_1: $\mu_1 \neq \mu_2$

构造统计量 $U = \dfrac{\bar{X} - \bar{Y}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

当H_0为真时, U~$N(0,1)$

|U|越大，对原假设越不利，故拒绝域应是形为{|U|>c}的小概率事件.

对给定显著性水平α， 检验拒绝域为

$$W = \left\{ |U| > u_{\frac{\alpha}{2}} \right\}$$

考虑两均值差 $\mu_1 - \mu_2$ 的区间估计, 则主元为

$$U = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

由$|U| \le u_{\frac{\alpha}{2}}$可得$\mu_1 - \mu_2$的区间估计。

2.当$\sigma_1^2=\sigma_2^2=\sigma^2$未知时， 检验$H_0$: $\mu_1=\mu_2$; H_1: $\mu_1\neq\mu_2$

构造统计量: $$T=\frac{\bar{X}-\bar{Y}}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

当H_0为真时, T~$t(n_1+n_2-2)$

|T|越大，对原假设越不利，故拒绝域应是形为{|T|>c}的小概率事件.

拒绝域为：$W=\left\{|T|>t_{\frac{\alpha}{2}}(n_1+n_2-2)\right\}$

考虑 $\mu_1-\mu_2$ 的区间估计, 则主元为

$$T=\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

$\sim t(n_1+n_2-2)$

由$|T|\le t_{\frac{\alpha}{2}}(n_1+n_2-2)$可得

$\mu_1-\mu_2$的区间估计.

- 三.基于成对数据的假设检验
- 设$(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$为独立同分布的二维正态总体样本, 令$Z_i=X_i-Y_i, i=1,2,\ldots n$。则$Z_i$=独立同分布，且服从$N(\mu,\sigma^2)$,
- 检验H_0: $\mu=0$; H_1: $\mu\neq0$
- 构造统计量: $T=\dfrac{\bar{Z}}{S_Z/\sqrt{n}}$，其中$S_Z$为$Z_i$的样本标准差。
- 当H_0为真时, $T\sim t(n-1)$。拒绝域为: $W=\left\{|T|>t_{\frac{\alpha}{2}}(n-1)\right\}$
- 注：与两总体均值差的检验不同处在哪里?

- 例8.2.1：比较两种安眠药的疗效，以10个失眠患者作为实验对象。对每个患者服药后，记录其延长的睡眠时间（单位：小时），数据如下：问在显著性水平α=0.05下，两种药品是否有显著差异？

A	1.9	0.8	1.1	0.1	-0.1	4.4	5.5	1.6	4.6	3.4
B	0.7	-1.6	-0.2	-1.2	-0.1	3.4	3.7	0.8	0	2.0

§ 8.3　正态总体方差的检验

以下均考虑显著性水平为α的假设检验。

一. 单个正态总体方差的检验

设总体X~N(μ,σ^2), $X_1,X_2,\ldots,X_n$是来自X的样本,

1. $H_0:\sigma^2=\sigma_0^2;H_1:\sigma^2\neq\sigma_0^2$ —χ^2检验法

考虑σ^2的区间估计，则主元为

$$\chi^2=\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$$

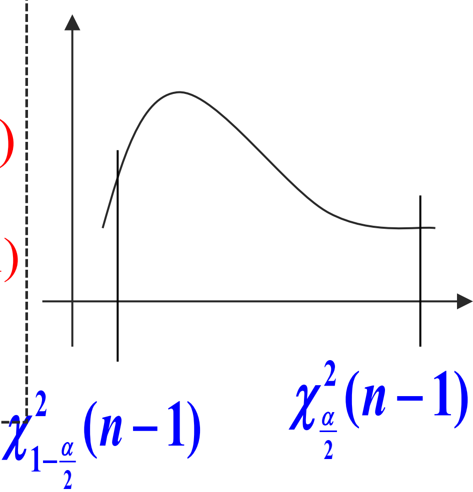
由$\chi^2_{1-\frac{\alpha}{2}}(n-1)\leq\chi^2\leq\chi^2_{\frac{\alpha}{2}}(n-1)$

可得σ^2的区间估计。

构造统计量 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$

当H_0为真时, $\chi^2\sim\chi^2(n\text{-}1)$

故检验拒绝域为



$$W=\left\{\chi^2<\chi^2_{1-\frac{\alpha}{2}}(n-1)\right\}\cup\left\{\chi^2>\chi^2_{\frac{\alpha}{2}}(n-1)\right\}$$

单个正态总体方差的单边检验

(1) $H_0: \sigma^2 \le \sigma_0^2; H_1: \sigma^2 > \sigma_0^2$

检验统计量： $$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$$

拒绝域为： $$W = \left\{\chi^2 > \chi_\alpha^2(n-1)\right\}$$

(2) $H_0: \sigma^2 \ge \sigma_0^2; H_1: \sigma^2 < \sigma_0^2$

检验统计量： $$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$$

拒绝域为： $W = \left\{\chi^2 < \chi_{1-\alpha}^2(n-1)\right\}$

例8.3.1 某厂生产的铜丝，质量一向比较稳定，今从中随机抽取10根检查其折断力，测得数据（单位：千克）如下： 575 576 570 569 572 582 577 580 572 585

设铜丝的折断力服从正态分布，试问在显著性水平0.05下是否可以相信该厂生产的铜丝折断力的方差为64？

解： $H_0:\sigma^2=64; H_1:\sigma^2\neq 64$

检验统计量： $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$

拒绝域为： $W=\left\{\chi^2<\chi^2_{1-\frac{\alpha}{2}}(9)\right\}\cup\left\{\chi^2>\chi^2_{\frac{\alpha}{2}}(9)\right\}$

$$=\left\{\chi^2<2.7\right\}\cup\left\{\chi^2>19.02\right\}$$

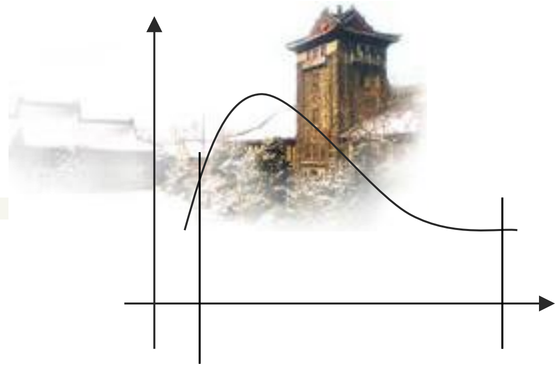
检验统计量的值为： $\chi^2=251.6/64\approx 3.93$

结论：因为2.7<3.93<19.02,所以不拒绝原假设，即认为该厂生产的铜丝折断力的方差为64。

二. 两个正态总体方差的检验

$X \sim N(\mu_1,\sigma_1^2)$, $\quad X_1,\cdots,X_{n_1}$

$Y \sim N(\mu_2,\sigma_2^2)$, $\quad Y_1,\cdots,Y_{n_2}$,两样本相互独立



检验 $H_0:\sigma_1^2=\sigma_2^2; H_1:\sigma_1^2\neq\sigma_2^2$

—F检验法，方差齐性检验

构造统计量 $F=\dfrac{S_1^2}{S_2^2}$

当H_0为真时, $F\sim F(n_1-1,n_2-1)$

检验拒绝域为

考虑$\dfrac{\sigma_1^2}{\sigma_2^2}$的区间估计，则主元为

$$F=\frac{S_1^2}{S_2^2}\cdot\frac{\sigma_2^2}{\sigma_1^2}\sim F(n_1-1,n_2-1)$$

由$F_{1-\frac{\alpha}{2}}\le F\le F_{\frac{\alpha}{2}}$可得区间估计。

$$W=\left\{F<F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)\right\}\cup\left\{F>F_{\frac{\alpha}{2}}(n_1-1,n_2-1)\right\}$$

$$=\left\{F<\frac{1}{F_{\frac{\alpha}{2}}(n_2-1,n_1-1)}\right\}\cup\left\{F>F_{\frac{\alpha}{2}}(n_1-1,n_2-1)\right\}$$

例8.3.3 甲乙两个农业试验区各分10个小区，各小区面积相等，作种植玉米的试验。甲区除施磷肥外，其它条件与乙区相同，试验结果玉米产量（单位：千克）如下：

甲： 62 57 65 60 63 58 57 60 60 58

乙： 56 59 56 57 58 57 60 55 57 55

设甲乙各小区的玉米产量均服从正态分布。

（1）在显著性水平0.1下，甲乙区玉米产量方差是否相同？

（2）在显著性水平0.05下，施磷肥对玉米产量有无影响？

解：设甲乙两区的产量分别服从分布 $N(\mu_1,\sigma_1^2)$和$N(\mu_2,\sigma_2^2)$

（1）方差的齐性检验　　$H_0:\sigma_1^2=\sigma_2^2;H_1:\sigma_1^2\neq\sigma_2^2$

检验统计量：　$F=\dfrac{S_1^2}{S_2^2}$

拒绝域为：

$$W=\left\{F<\frac{1}{F_{\frac{\alpha}{2}}(n_2-1,n_1-1)}\right\}\cup\left\{F>F_{\frac{\alpha}{2}}(n_1-1,n_2-1)\right\}$$

$$=\left\{F<\frac{1}{3.18}\right\}\cup\{F>3.18\}$$

统计量的值为：　$f=64/24\approx2.667$

结论：　因为1/3.18<2.667<3.18,所以不拒绝原假设，
即认为方差相同。

（2）均值的比较 $H_0: \mu_1 = \mu_2; H_1: \mu_1 \neq \mu_2$

检验统计量： $$T = \frac{\bar{X} - \bar{Y}}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

拒绝域为： $$W = \left\{ |T| > t_{\frac{\alpha}{2}}(n_1 + n_2 - 2) \right\}$$

$$= \left\{ |T| > t_{0.025}(18) = 2.1 \right\}$$

统计量的值为： $t \approx 3.034$

结论： 因为3.034>2.1，所以拒绝原假设，即认为施磷肥对玉米产量有影响。

关于假设检验与区间估计的关系

例： 单个正态总体均值的假设检验和区间估计

设总体$X\sim N(\mu,\sigma^2)$, $X_1,X_2,\ldots,X_n$是来自X的样本，$(\sigma^2$未知$)$

1. 检验$H_0:\mu=\mu_0$; $H_1:\mu\neq\mu_0$

统计量： $$T=\frac{\bar{X}-\mu_0}{S/\sqrt{n}}$$

拒绝域为$W=\left\{|T|>t_{\frac{\alpha}{2}}(n-1)\right\}$

检验水平为α的接受域为

$$\overline{W}=\left\{\left|\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\right|\le t_{\frac{\alpha}{2}}(n-1)\right\}$$

$$=\{\overline{X}-\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}\le\mu_0\le\overline{X}+\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}\}$$

2. 考虑总体均值μ的点估计，则主元为

$$T=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$$

由$\left|\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right|\le u_{\frac{\alpha}{2}}$可得

置信度为$1-\alpha$的置信区间为

$$[\overline{X}-\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}},\overline{X}+\frac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}]$$

§ 8.4 拟合优度检验

- 总体$X\sim F(x)$类型未知，$F_0(x)$是已知的可能含有未知参数的分布函数。$X_1,\ldots,X_n$是样本，$F_0(x)$称为理论分布。
- 假设检验：$H_0:F(x)=F_0(x)$； $H_1:F(x)\neq F_0(x)$。
- 或更一般的：

$$H_0:F(x)=F_0(x;\theta);H_1:F(x)\neq F_0(x;\theta)$$

- 称为拟合优度检验。

- 先考虑最简单的情况：

$$H_0: P(X=x_i)=p_i, i=1,2,\cdots,k.$$

- 1.计算样本$X_1,\ldots,X_n$中x_i的频数n_i，i=1,2,…,k;
- 2.计算实际频数与理论频数的差距，构造皮尔逊检验统计量：

$$\chi^2=\sum_{i=1}^{k}\frac{(n_i-np_i)^2}{np_i}$$

- 3.因为χ^2的极限分布为$\chi^2(k-1)$，于是拒绝域为

$$W=\{\chi^2>\chi^2_\alpha(k-1)\}$$

- 再考虑复杂的情况：

$$H_0 : P(X = x_i) = p_i(\theta), i = 1, 2, \cdots, k . \theta \in R^r$$

- 1.计算样本$X_1,\ldots,X_n$中x_i的频数n_i，i=1,2,…,k;
- 2.计算实际频数与理论频数的差距，构造皮尔逊检验统计量：

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i(\hat{\theta}))^2}{np_i(\hat{\theta})}$$，其中$\hat{\theta}$为θ的极大似然估计。

- 3.因为χ^2的极限分布为χ^2(k-r-1)，于是拒绝域为

$$W = \{\chi^2 > \chi^2_\alpha(k - r - 1)\}$$

- 下面介绍一般χ^2拟合检验法的基本想法：

$$H_0: F(x) = F_0(x); H_1: F(x) \neq F_0(x) \text{或}$$
$$H_0: F(x) = F_0(x;\theta); H_1: F(x) \neq F_0(x;\theta)$$

- 1.把随机试验的样本空间分成k个互不相容的事件$A_1, A_2, \ldots, A_k$，也称为k个组；这样就可以通过每一组上概率是否相等来检验原假设了（离散化）。
- 2.计算样本$X_1, \ldots, X_n$落在每一组上的实际频数n_i;，i=1,2,…,k;

- **3.计算落在每一组上的理论频数$n\ p_i$,其中p_i是当H_0为真时，落在每一组上的概率；若分布族存在参数θ，则计算参数θ的极大似然估计，得到p_i的估计如下：**

$$\hat{p}_i = P(X \in A_i \mid \theta = \hat{\theta}),\text{其中}\hat{\theta}\text{为}\theta\text{的极大似然估计。}$$

- **4.当H_0为真时，计算皮尔逊拟合检验统计量**

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}\text{或}$$

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - n\hat{p}_i)^2}{n\hat{p}_i},\text{其中}\hat{p}_i = P(X \in A_i \mid \theta = \hat{\theta}),\hat{\theta}\text{为}\theta\text{的极大似然估计。}$$

- 上述构造出的统计量，取值越大，越应该拒绝原假设。但要确定一个比较大的临界值，就需要知道统计量的分布。
- **定理：（Pearson-Fisher定理）**在H_0为真时，不论总体服从什么分布，当n很大时，上述定义的统计量χ^2服从自由度为k-r-1的χ^2分布，其中r是$F_0(x)$中未知参数的个数。即

$$\chi^2=\sum_{i=1}^{k}\frac{(n_i-n\hat{p}_i)^2}{n\hat{p}_i}\xrightarrow{d}\chi^2(k-r-1)$$

- 5.给定显著性水平α，拒绝域为

$$W=\{\chi^2>\chi_\alpha^{\,2}(k-r-1)\}$$

- 注：1.上述方法称为χ^2拟合检验法。它只适合于大样本情况，一般要求样本容量不低于50。
- 2.每一组上的$n\,p_i$不能太小，一般不少于5；如果小于5，可和临近的组进行合并。分组数k是一般为$n/5$—$n/10$之间，当n很大时，一般取20组即可。
- 3.分布中如果含有未知参数，则计算统计量值之前要先对其进行估计，一般采用极大似然估计。

- 例8.4.1：假设某企业早、中、晚三班在发生事故的概率上无差别，均为1/3。一段时间记录到15次生产事故，其中早、晚各6次，中班3次。这些数据使人怀疑各班次的事故率不完全相同，需要进一步检验早、中、晚三班在发生事故的概率。
- 解：

$$\chi^2 = \sum_{i=1}^{3} \frac{(n_i - np_i)^2}{np_i} = 1.2 < \chi_\alpha^2(3-1) = \chi_{0.05}^2(2) = 5.991$$

- 例8.4.2：在一实验中，每隔一定时间观察一次由某种铀所放射的到达计数器的α粒子数，共观察了100次，得结果如下：

x_i	0	1	2	3	4	5	6	7	8	9	10	11	≥12
n_i	1	5	16	17	26	11	9	9	2	1	2	1	0

- 其中n_i是观测到有x_i个粒子的频数。试问到达计数器的α粒子数是否服从泊松分布？
- 解：设X表示到达计数器的α粒子数，则需检验H_0：$X\sim P(\lambda)$。

- （1）对样本空间进行分组。因为X取值是离散的，初分组时可以考虑一个取值一组，因此先分为13组。适当合并后为8组。
- （2）每一组的实际频数已经在表里列出。
- （3）计算每一组的理论频数。显然理论概率中含有未知参数，因此我们需要对其进行估计。而参数λ显然就是总体均值，因此它的极大似然估计值为：

$$\hat{\lambda}_{ML} = \bar{X} = \frac{\sum_{i=0}^{11} x_i n_i}{n} = 4.2$$

- 结果如下表：

x_i	n_i	$\hat{p}_i$	$n\hat{p}_i$	$n_i - n\hat{p}_i$	$\frac{(n_i - n\hat{p}_i)^2}{n\hat{p}_i}$
0,1	6	0.078	7.8	-1.8	0.415
2	16	0.132	13.2	2.8	0.594
3	17	0.185	18.5	-1.5	0.122
4	26	0.194	19.4	6.6	2.245
5	11	0.163	16.3	-5.3	1.723
6	9	0.114	11.4	-2.4	0.505
7	9	0.069	6.9	2.1	0.639
≥ 8	6	0.065	6.5	-0.5	0.0385

- （4）拒绝域为：

$$\chi^2=\sum_{i=1}^{8}\frac{(n_i-n\hat{p}_i)^2}{n\hat{p}_i}>\chi^2_\alpha(8-1-1)=\chi^2_{0.05}(6)=12.592$$

- （5）计算$\chi^2=6.2815<12.592$，所以不拒绝原假设，即认为到达计数器的α粒子数服从泊松分布。

- 例8.4.3：研究某种建筑材料的抗压强度是否服从正态分布，200次试验的数据如下：

压强区间（公斤/厘米）	实际频数
190-200	10
200-210	26
210-220	56
220-230	64
230-240	30
240-250	14

- 试在显著性水平α=0.05下，检验H_0：X服从正态分布。

§8.5 独立性检验

- 一. 独立性检验

1. 吸烟和患肺癌有关吗？

2. 学习成绩好坏和性别有关吗？

3. 能都找到好工作和政治面貌有关吗？

这些都是要讨论两个变量之间是否独立的问题。

- 设$(X_1,Y_1),\cdots,(X_n,Y_n)$是来自二维总体$(X,Y)$的一个样本，独立性检验就是要检验假设：

$$H_0: X\text{与}Y\text{独立}; \ H_1: X\text{与}Y\text{不独立}$$

- 我们把X的取值分成r个部分，分别用$A_1, \cdots, A_r$表示；将Y的取值也分为s个部分，用$B_1,\cdots,B_s$表示。分布列如下：

$X \backslash Y$	B_1	…	B_s	
A_1	p_{11}	…	p_{1s}	$p_{1.}$
…	…	…	…	…
A_r	p_{r1}	…	p_{rs}	$p_{r.}$
	$p_{.1}$	…	$p_{.s}$	1

- 因此原假设H_0:X与Y独立转化为

$$H_0: p_{ij} = p_{i\cdot}p_{\cdot j},\quad i,=1,\cdots,r, j=1,\cdots,s.$$

- 上述检验相当于离散情形下带参数的拟合优度检验。**问题：参数是什么？k=?，r=?。**
- 假设以n_{ij}表示落在第A_iB_j组内的实际频数。

$$n_{i\cdot}=\sum_{j=1}^{s}n_{ij},\ n_{\cdot j}=\sum_{i=1}^{r}n_{ij},\ n=\sum_{i=1}^{r}\sum_{j=1}^{s}n_{ij}$$

- 列联表如下：

X \ Y	B_1	…	B_s	行和
A_1	n_{11}	…	n_{1s}	$n_{1.}$
…	…	…	…	…
A_r	n_{r1}	…	n_{rs}	$n_{r.}$
列和	$n_{.1}$	…	$n_{.s}$	n

- 先求参数的极大似然估计。$(X_1,Y_1),\cdots,(X_n,Y_n)$的似然函数为：

$$L=\prod_{i=1}^{r}\prod_{j=1}^{s}p_{ij}^{n_{ij}}$$

- 当原假设成立时：

$$L=\prod_{i=1}^{r}\prod_{j=1}^{s}(p_{i\cdot}p_{\cdot j})^{n_{ij}}=\prod_{i=1}^{r}p_{i\cdot}^{n_{i\cdot}}\prod_{j=1}^{s}p_{\cdot j}^{n_{\cdot j}}$$

$$=\left(1-\sum_{i=1}^{r-1}p_{i\cdot}\right)^{n_{r\cdot}}\left(1-\sum_{j=1}^{s-1}p_{\cdot j}\right)^{n_{\cdot s}}\prod_{i=1}^{r-1}p_{i\cdot}^{n_{i\cdot}}\prod_{j=1}^{s-1}p_{\cdot j}^{n_{\cdot j}}$$

- 对数似然函数为：

$$\ln L = n_{r\cdot}\ln\left(1-\sum_{i=1}^{r-1}p_{i\cdot}\right)+n_{\cdot s}\ln\left(1-\sum_{j=1}^{s-1}p_{\cdot j}\right)+\sum_{i=1}^{r-1}n_{i\cdot}\ln p_{i\cdot}+\sum_{j=1}^{s-1}n_{\cdot j}\ln p_{\cdot j}$$

- 令

$$\begin{cases}\dfrac{\partial \ln L}{\partial p_{i\cdot}}=\dfrac{-n_{r\cdot}}{1-\sum\limits_{i=1}^{r-1}p_{i\cdot}}+\dfrac{n_{i\cdot}}{p_{i\cdot}}=0 & i=1,\cdots,r-1\\[2ex] \dfrac{\partial \ln L}{\partial p_{\cdot j}}=\dfrac{-n_{\cdot s}}{1-\sum\limits_{j=1}^{s-1}p_{\cdot j}}+\dfrac{n_{\cdot j}}{p_{\cdot j}}=0 & j=1,\cdots,s-1\end{cases}$$

- 解得：

$$\begin{cases} p_{i\cdot} = \dfrac{n_{i\cdot}}{n} & i = 1, \cdots, r-1 \\ p_{\cdot j} = \dfrac{n_{\cdot j}}{n} & j = 1, \cdots, s-1 \end{cases}$$

- 因此可以得到皮尔逊χ^2检验统计量如下：

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{s} \frac{(n_{ij} - n\hat{p}_{i\cdot}\cdot\hat{p}_{\cdot j})^2}{n\hat{p}_{i\cdot}\cdot\hat{p}_{\cdot j}} = \sum_{i=1}^{r}\sum_{j=1}^{s} \frac{(n_{ij} - n_{i\cdot}\cdot n_{\cdot j}/n)^2}{n_{i\cdot}\cdot n_{\cdot j}/n}$$

$$\rightarrow \chi^2(rs-(r+s-2)-1) = \chi^2((r-1)(s-1))$$

- 于是拒绝域为：$W = \{\chi^2 > \chi^2_\alpha((r-1)(s-1))\}$

- 例8.5.1：中央民族大学教育学研究室的课题“学习成绩与道德的认识水平之间的关系”要研究的问题是，学生的学习成绩与道德认识水平有关系吗？他们将学生的学习成绩分为优、良、中、差四个等级，将道德认识水平分为好、中上、中下、差四档，随机调查了150名同学，调查结果如下表：**问：从调查数据看，道德认识水平与学习成绩有没有关系？**

B道德 / A学习	B1	B2	B3	B4	
A1	20	8	1	0	29
A2	5	40	14	1	60
A3	0	2	18	6	26
A4	0	1	11	23	35
	25	51	44	30	n＝150

求解过程

- 解： H_0：学习成绩与道德的认识水平之间没有关系，即 H_0：X与Y独立。进一步转化为

$$H_0: p_{ij} = p_{i\cdot}p_{\cdot j}, \quad i, = 1, \cdots, 4, j = 1, \cdots, 4.$$

- 检验统计量：

$$\chi^2 = \sum_{i=1}^{4}\sum_{j=1}^{4}\frac{(n_{ij} - n_{i\cdot}n_{\cdot j}/n)^2}{n_{i\cdot}n_{\cdot j}/n}$$

- 拒绝域： $\chi^2 > \chi^2_\alpha(9) = \chi^2_{0.05}(9) = 16.92$
- 经计算可知χ^2=168.8174>16.92，所以拒绝原假设。即认为学习成绩与道德的认识水平之间存在关系。

二. 齐一性检验

有r个总体， 取值为1,2,…s个水平。第i个总体分布为

$$P(X_i = j) = f_j(\theta_i), j = 1, 2, \cdots, s; i = 1, 2, \cdots, r. \theta_i \in R^q$$

下面检验这r个总体分布是否相同。即

$$H_0 : \theta_1 = \theta_2 = \cdots = \theta_r = \theta$$

建表如下（假设以*nij*表示落在第*i*行*j*列内的实际频数）：

水平 总体	1	…	s	
1	n_{11}	…	n_{1s}	$n_{1.}$
…	…	…	…	…
r	n_{r1}	…	n_{rs}	$n_{r.}$
	$n_{.1}$	…	$n_{.s}$	n

利用样本对 θ_i及公共值θ做极大似然估计$\hat{\theta}_i, \hat{\theta}$。

可以构造如下χ**2**分布的统计量：

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(n_{i\cdot}f_j(\hat{\theta}_i) - n_{i\cdot}f_j(\hat{\theta}))^2}{n_{i\cdot}f_j(\hat{\theta})}$$

$$\xrightarrow{d} \chi^2((r-1)q)$$

拒绝域为：

$$W = \{\chi^2 > \chi_\alpha^2((r-1)q)\}$$

下面考虑特殊情形：

$$P(X_i = j) = p_{ij}, j = 1,2,\cdots,s; i = 1,2,\cdots,r.$$

也就是说，这r个总体的分布完全未知。

此时q=s-1。假设检验为：

$$H_0 : p_{1j} = p_{2j} = \cdots = p_{rj} = p_j, j = 1,2,\cdots,s.$$

不难计算$\hat{p}_{ij} = \dfrac{n_{ij}}{n_{i\cdot}}, \hat{p}_j = \dfrac{n_{\cdot j}}{n}.$

因此χ^2检验统计量为$\chi^2 = \displaystyle\sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(n_{i\cdot}\hat{p}_{ij} - n_{i\cdot}\hat{p}_j)^2}{n_{i\cdot}\hat{p}_j}$

$$= \sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(n_{ij} - n_{i\cdot}n_{\cdot j}/n)^2}{n_{i\cdot}n_{\cdot j}/n} \xrightarrow{d} \chi^2((r-1)(s-1))$$

例题：习题最后一题。