天善智能
TIANSHAN SOFT
专注于商业智能BI和大数据的垂直社区平台

常见参数假设检验

Allen

www.hellobi.com

课程目录

- u检验
- t检验
- 卡方检验
- 小结

天善智能
TIANSHAN SOFT

u检验

- 设$X_1,X_2,\ldots,X_n$是取自正态母体$N(\mu,\sigma^2)$的一个子样，其中$\sigma^2=\sigma_0^2$为已知常数，要检验假设$H_0:\mu=\mu_0$ ，$H_1:\mu\neq\mu_0$

给出原假设和备择假设

- 如果原假设$H_0:\mu=\mu_0$为真，那么子样均值$\overline{X}$应当在μ_0上下随机摆动，而不会偏离μ_0太大，所以其临界域的结构形如$\left(\left|\overline{X}-\mu_0\right|\geq k\right)$
- 为了便于查找分位数表，将统计量改写成$u=\dfrac{\overline{X}-\mu_0}{\sigma_0/\sqrt{n}}$，在原假设为真是服从标准正态分布

给出检验统计量

u检验

- 给定显著性水平α，当原假设H_0为真时，$P\left(|u| \geq u_{1-\frac{\alpha}{2}}\right)=\alpha$，这里的$u_{1-\frac{\alpha}{2}}$是标准正态分布$N(0,1)$的$1-\frac{\alpha}{2}$分位点，查表就可得到，进而求出临界域

$$C=\left(|u| \geq u_{1-\frac{\alpha}{2}}\right)$$

给出显著性水平并求出临界域

- 根据子样观测值算出u的值，若$|u| \geq u_{1-\frac{\alpha}{2}}$则拒绝原假设$H_0: \mu=\mu_0$，并认为母体均值与原假设有显著差异

以上这种方法称为u检验

天善智能
TIANSHAN SOFT

t检验

- 设$X_1, X_2, \ldots, X_n$是取自正态母体$N(\mu, \sigma^2)$的一个子样，其中$\sigma^2=\sigma_0^2$为未知常数，要检验假设$H_0: \mu = \mu_0$，$H_1: \mu \neq \mu_0$

给出原假设和备择假设

- 之前的u检验中的统计量$u=\frac{\overline{X}-\mu_0}{\sigma_0/\sqrt{n}}$中的$\sigma_0$未知，固不能使用u检验，并且首先要选取一个不含未知参数的σ^2的统计量。

- 选取方差的无偏估计量$S^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$去代替母体方差$\sigma^2$，这样就得到t统计量$t=\frac{\overline{X}-\mu_0}{S^*}\sqrt{n}$，这个统计量服从自由度为$n-1$的t分布

给出检验统计量

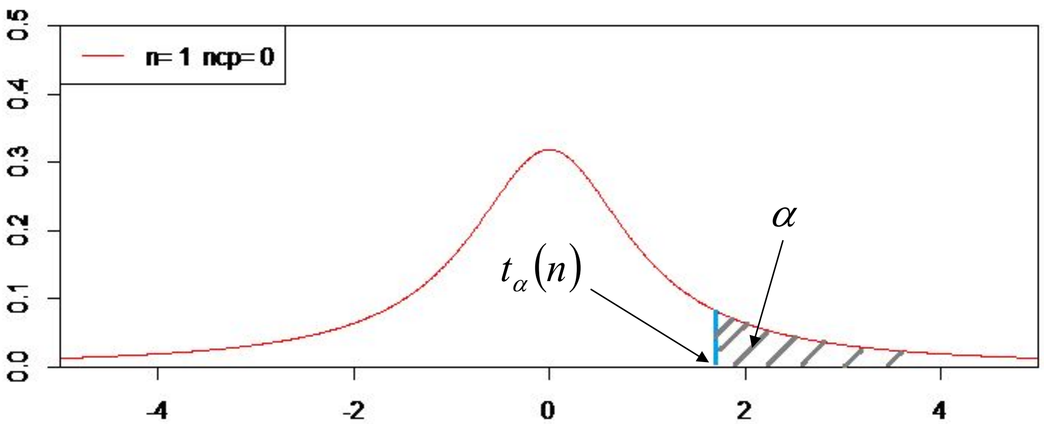
天善智能
TIANSHAN SOFT

t检验

- 给定显著性水平α ，当原假设H_0为真时，$P\left(|t|\ge t_{1-\frac{\alpha}{2}}(n-1)\right)=\alpha$，这里的$t_{1-\frac{\alpha}{2}}(n-1)$是自由度为$n-1$的t分布的$1-\frac{\alpha}{2}$分位点，查表就可得到，进而求出临界域

$$C=\left(|t|\ge t_{1-\frac{\alpha}{2}}(n-1)\right)$$

给出显著性水平并求出临界域



t检验

- 根据子样观测值算出t的值，若$|t| \geq t_{1-\frac{\alpha}{2}}(n-1)$则拒绝原假设 $H_0 : \mu = \mu_0$ ，并认为母体均值与原假设有显著差异

以上这种方法称为t检验

卡方检验——母体均值已知

- 设$X_1, X_2, \ldots, X_n$是取自正态母体$N(\mu, \sigma^2)$的一个子样，要检验假设$H_0: \sigma^2 = \sigma_0{}^2$，

$H_1: \sigma^2 \neq \sigma_0{}^2$

给出原假设和备择假设

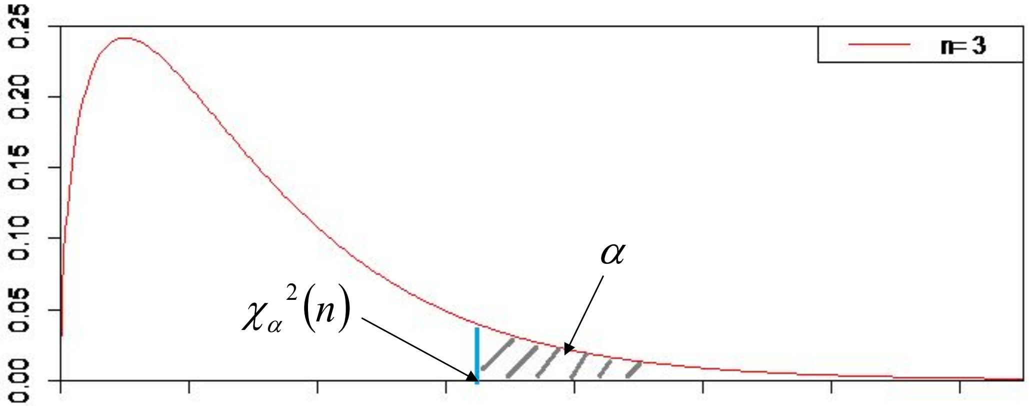
- 下面要分别对μ已知和μ未知两种情况进行讨论

- $\mu = \mu_0$为已知常数构造统计量$\chi^2 = \dfrac{\sum_{i=1}^{n}(X_i - \mu_0)^2}{\sigma_0{}^2}$服从自由度为n的卡方分布

给出检验统计量

天善智能 TIANSHAN SOFT

卡方检验——母体均值已知



- 给定显著性水平 α ，由图可看出卡方分布的临界域的格式为 $\{\chi^2 \le k_1\} \cup \{\chi^2 \ge k_2\}$ 概率表示为 $P(\chi^2 \le k_1) = \alpha_1$ 和 $P(\chi^2 \ge k_2) = \alpha_2$ 其中 $\alpha_1 + \alpha_2 = \alpha$ ，显然 k_1 和 k_2 分别是自由度为n的卡方分布的 α_1 和 $1-\alpha_2$ 分位点

卡方检验——母体均值已知

- 若取 α_1 和 α_2 分别为 $\alpha/2$，这时 k_1 和 k_2 就分别是卡方分布的 $\frac{\alpha}{2}$ 和 $1-\frac{\alpha}{2}$ 分位点，即 $k_1=\chi_{\alpha/2}{}^2(n)$，$k_2=\chi_{1-\alpha/2}{}^2(n)$ 这时得到的临界域为

$$C=\{\chi^2\le\chi_{\alpha/2}{}^2(n)\}\cup\{\chi^2\ge\chi_{1-\alpha/2}{}^2(n)\}$$

给出显著性水平并求出临界域

- 根据子样观测值算出检验统计量值，若小于 $\chi_{\alpha/2}{}^2(n)$ 或者大于 $\chi_{1-\alpha/2}{}^2(n)$ 则拒绝原假设，并认为母体均值与原假设有显著差异

以上这种方法称为卡方检验

卡方检验——母体均值未知

- 设$X_1, X_2, \ldots, X_n$是取自正态母体$N(\mu, \sigma^2)$的一个子样，要检验假设$H_0: \sigma^2 = {\sigma_0}^2$，

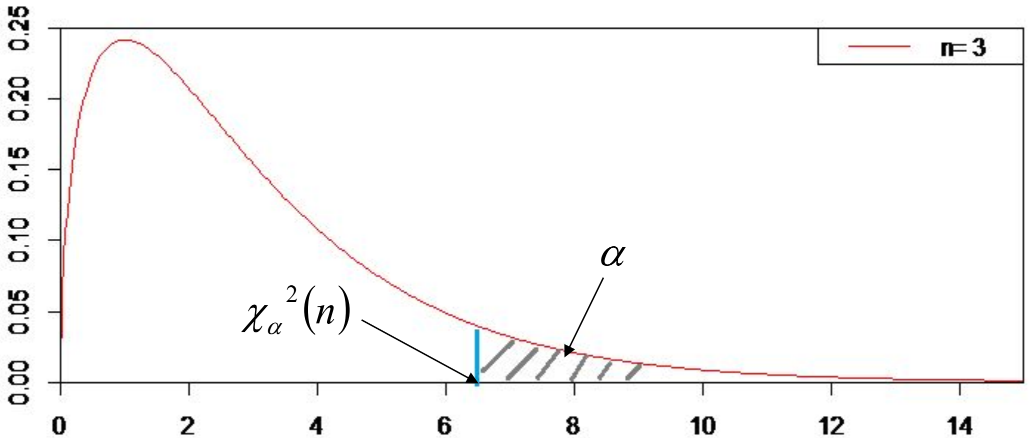
$H_1: \sigma^2 \neq {\sigma_0}^2$

给出原假设和备择假设

- μ为未知，则用子样均值$\overline{X}$去代替母体均值构造统计量$\chi^2 = \dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{{\sigma_0}^2}$ 服从自由度为$n-1$ 的卡方分布

给出检验统计量

卡方检验——母体均值未知



- 给定显著性水平α ，由图可看出卡方分布的临界域的格式为

$\{\chi^2 \le k_1\} \cup \{\chi^2 \ge k_2\}$概率表示为$P(\chi^2 \le k_1) = \alpha_1$ 和 $P(\chi^2 \ge k_2) = \alpha_2$ 其中 $\alpha_1 + \alpha_2 = \alpha$ ，显然k_1

和k_2分别是自由度为n的卡方分布的α_1和$1-\alpha_2$分位点

卡方检验——母体均值未知

- 若取α_1和α_2分别为$\alpha/2$，这时k_1和k_2就分别是卡方分布的$\frac{\alpha}{2}$和$1-\frac{\alpha}{2}$分位点，即$k_1=\chi_{\alpha/2}{}^2(n-1)$，$k_2=\chi_{1-\alpha/2}{}^2(n-1)$这时得到的临界域为

$$C=\left\{\chi^2\le\chi_{\alpha/2}{}^2(n-1)\right\}\cup\left\{\chi^2\ge\chi_{1-\alpha/2}{}^2(n-1)\right\}$$

给出显著性水平并求出临界域

- 根据子样观测值算出检验统计量值，若小于$\chi_{\alpha/2}{}^2(n-1)$或者大于$\chi_{1-\alpha/2}{}^2(n-1)$则拒绝原假设，并认为母体均值与原假设有显著差异

以上这种方法称为卡方检验

天善智能 TIANSHAN SOFT

小结

- u检验
- t检验
- 卡方检验
- 小结