假设检验

续本达

清华大学 工程物理系

2024-11-20

复习

区间估计的步骤

定义 (枢轴量)

仅含一个待估参数的样本的连续函数，且分布不依赖于未知参数。

1. 选取枢轴量
2. 由分位点定义建立不等式
3. 解出不等式

正态总体

	待估参数	其他参数	枢轴量及其分布	置信区间	单侧置信限
一个正态总体	μ	σ^2 已知	$Z=\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\left(\bar{X}\pm\dfrac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}\right)$	$\bar{\mu}=\bar{X}+\dfrac{\sigma}{\sqrt{n}}z_\alpha$ $\underline{\mu}=\bar{X}-\dfrac{\sigma}{\sqrt{n}}z_\alpha$
	μ	σ^2 未知	$t=\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$	$\left(\bar{X}\pm\dfrac{S}{\sqrt{n}}t_{\frac{\alpha}{2}}(n-1)\right)$	$\bar{\mu}=\bar{X}+\dfrac{S}{\sqrt{n}}t_\alpha(n-1)$ $\underline{\mu}=\bar{X}-\dfrac{S}{\sqrt{n}}t_\alpha(n-1)$
	σ^2	μ 已知	$Q=\sum_{i=1}^n\left(\dfrac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n)$	$\left(\dfrac{\sum_{i=1}^n(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)},\dfrac{\sum_{i=1}^n(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}\right)$	$\overline{\sigma^2}=\dfrac{\sum_{i=1}^n(X_i-\mu)^2}{\chi^2_{1-\alpha}(n)}$ $\underline{\sigma^2}=\dfrac{\sum_{i=1}^n(X_i-\mu)^2}{\chi^2_\alpha(n)}$
	σ^2	μ 未知	$\chi^2=\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$	$\left(\dfrac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)},\dfrac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$	$\overline{\sigma^2}=\dfrac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)}$ $\underline{\sigma^2}=\dfrac{(n-1)S^2}{\chi^2_\alpha(n-1)}$

思考：连续参数和样本的连续函数

1. 不管是连续还是离散分布，它们的参数都是连续的。
   - 泊松分布离散，但是参数 λ 连续
   - 二项分布离散，但是参数 p 连续
2. 统计量的定义：样本 $(X_1, X_2, \cdots, X_n)$ 的不含有未知参数的 连续函数 $g(X_1, X_2, \cdots, X_n)$。
   - 枢轴量：仅含一个待估参数的样本的 连续函数
3. 离散的参数的例子
   - 对、错
   - 有、无

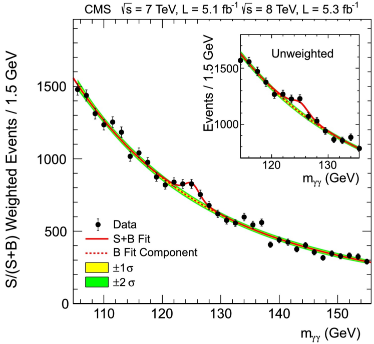
思考：连续参数和样本的连续函数

1. 不管是连续还是离散分布，它们的参数都是连续的。
   - 泊松分布离散，但是参数 λ 连续
   - 二项分布离散，但是参数 p 连续
2. 统计量的定义：样本 $(X_1, X_2, \cdots, X_n)$ 的不含有未知参数的 连续函数 $g(X_1, X_2, \cdots, X_n)$。
   - 枢轴量：仅含一个待估参数的样本的 连续函数
3. 离散的参数的例子
   - 对、错
   - 有、无

假设检验

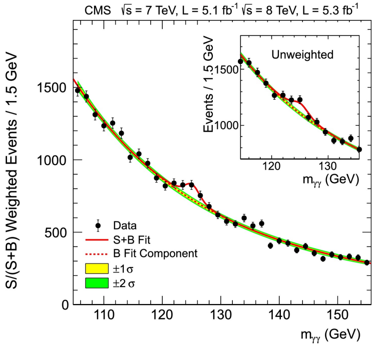
Higgs 玻色子的发现

- Higgs 玻色子是电弱统一理论预言的基本粒子
- 通过 Higgs 机制赋予其它基本粒子质量
- $H \to \gamma\gamma$ 为 2012 Higgs 发现的反应
- 从图中的看，我们有多大信心认为 Higgs 被发现?



Higgs 玻色子的发现

- Higgs 玻色子是电弱统一理论预言的基本粒子
- 通过 Higgs 机制赋予其它基本粒子质量
- $H \to \gamma\gamma$ 为 2012 Higgs 发现的反应
- 从图中的看，我们有多大信心认为 Higgs 被发现？



引例

例

口罩生产厂检验规定：次品率 p 不超过 4 % 才能出厂．现从一万件产品中任意抽查 12 件发现 3 件次品，问该批口罩能否出厂？

假设 $p \leq 0.04$，12 件产品中发现 3 件次品的概率为

$$\binom{12}{3} p^3 (1-p)^9 = 0.0097 < 0.01$$

事出反常必有妖

这是 小概率事件 ，一般在一次试验中不会发生，现一次试验竟然发生，故认为原假设不成立，即该批产品次品率 $p > 0.04$，因此该批产品不能出厂。

引例

例

口罩生产厂检验规定：次品率 p 不超过 4% 才能出厂. 现从一万件产品中任意抽查 12 件发现 3 件次品, 问该批口罩能否出厂?

假设 $p \leq 0.04$，12 件产品中发现 3 件次品的概率为

$$\binom{12}{3} p^3 (1-p)^9 = 0.0097 < 0.01$$

事出反常必有妖

这是 小概率事件 ，一般在一次试验中不会发生，现一次试验竟然发生，故认为原假设不成立，即该批产品次品率 $p > 0.04$，因此该批产品不能出厂。

引例

例

口罩生产厂检验规定：次品率 p 不超过 4 % 才能出厂．现从一万件产品中任意抽查 12 件发现 3 件次品，问该批口罩能否出厂？

假设 $p \le 0.04$，12 件产品中发现 3 件次品的概率为

$$\binom{12}{3} p^3 (1-p)^9 = 0.0097 < 0.01$$

事出反常必有妖

这是 小概率事件 ，一般在一次试验中不会发生，现一次试验竟然发生，故认为原假设不成立，即该批产品次品率 $p > 0.04$，因此该批产品不能出厂。

概念

1. 事先对总体参数的值作出某种假设
2. 然后利用样本信息来检验假设是否成立
   例如检验产品是否达到出厂标准

小概率原理

- 概率很小的事件在 一次 试验中是不会发生的；在一次试验中小概率事件如果发生了，我们就有理由认为提出的原假设是不对的。
- 小概率由研究者事先确定，一般取 1 %，5 % 或 10 % 。小概率的取值不同，假设检验的结果可能不同。在统计学中，小概率又叫 显著性水平 ，因此，假设检验又称为 显著性检验 。

概念

1. 事先对总体参数的值作出某种假设
2. 然后利用样本信息来检验假设是否成立
   例如检验产品是否达到出厂标准

小概率原理

- 概率很小的事件在 一次 试验中是不会发生的；在一次试验中小概率事件如果发生了，我们就有理由认为提出的原假设是不对的。
- 小概率由研究者事先确定，一般取 1 %，5 % 或 10 % 。小概率的取值不同，假设检验的结果可能不同。在统计学中，小概率又叫 显著性水平 ，因此，假设检验又称为 显著性检验 。

步骤

思路

1. 对所研究的总体做某个假设。
   - 假设其未知的 $\mathrm{E}(X) = 90$；
   - 假设两个总体的 $\mathrm{E}(X) = \mathrm{E}(Y)$ 等。
2. 通过抽样的样本值来检验是否接受假设
   - 是否接受的标准：看在假设条件下发生抽样结果的事件是否为小概率事件。
   - 若是，就拒绝假设；若不是，就接受假设。

引例

工厂生产螺丝钉，标准强度为 68。实际生产的强度 X 服从 $N(\mu, 3.62^2)$。若 $\mathrm{E}(X) = \mu = 68$, 则认为这批螺钉符合要求, 否则认为不符合要求. 现从整批螺钉中取容量为 36 的样本, 其均值为 $\overline{X} = 68.5$。符合要求吗?

思路

1. 对所研究的总体做某个假设。
   - 假设其未知的 $\mathrm{E}(X) = 90$；
   - 假设两个总体的 $\mathrm{E}(X) = \mathrm{E}(Y)$ 等。
2. 通过抽样的样本值来检验是否接受假设
   - 是否接受的标准：看在假设条件下发生抽样结果的事件是否为小概率事件。
   - 若是，就拒绝假设；若不是，就接受假设。

引例

工厂生产螺丝钉，标准强度为 68。实际生产的强度 X 服从 $N(\mu, 3.62^2)$。若 $\mathrm{E}(X) = \mu = 68$, 则认为这批螺钉符合要求, 否则认为不符合要求. 现从整批螺钉中取容量为 36 的样本, 其均值为 $\overline{X} = 68.5$。符合要求吗?

思路

1. 对所研究的总体做某个假设。
   - 假设其未知的 $\mathrm{E}(X) = 90$；
   - 假设两个总体的 $\mathrm{E}(X) = \mathrm{E}(Y)$ 等。
2. 通过抽样的样本值来检验是否接受假设
   - 是否接受的标准：看在假设条件下发生抽样结果的事件是否为小概率事件。
   - 若是，就拒绝假设；若不是，就接受假设。

引例

工厂生产螺丝钉，标准强度为 68。实际生产的强度 X 服从 $N(\mu, 3.62^2)$。若 $\mathrm{E}(X) = \mu = 68$, 则认为这批螺钉符合要求, 否则认为不符合要求. 现从整批螺钉中取容量为 36 的样本, 其均值为 $\overline{X} = 68.5$。符合要求吗?

原假设 null hypothesis 与 备择假设 alternative hypothesis

原假设定义

待检验的假设，又称“零假设”。

1. 总是有等号：$=, \leq, \geq$
2. 记为 H_0
   - $H_0 \colon \mu = 68$

备择假设定义

在原假设被拒绝后可供选择的假设

1. 总是有不等号：$<, >$
2. 记为 H_1
   - $H_1 \colon \mu \neq 68$

原假设 null hypothesis 与 备择假设 alternative hypothesis

原假设定义

待检验的假设，又称“零假设”。

1. 总是有等号： $=, \leq, \geq$
2. 记为 H_0
   - H_0:$\mu = 68$

备择假设定义

在原假设被拒绝后可供选择的假设

1. 总是有不等号: $<, >$
2. 记为 H_1
   - H_1:$\mu \neq 68$

选择检验统计量

检验统计量定义

用于假设检验决策的统计量

- 若原假设正确，则 $|\bar{x}-\mu|$ 不应太大，且

$$Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}=\frac{\overline{X}-68}{3.6/\sqrt{36}}\sim N(0,1)$$

- 确定检验统计量为 Z

选择检验统计量

检验统计量定义

用于假设检验决策的统计量

- 若原假设正确，则 $|\overline{x} - \mu|$ 不应太大，且

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\overline{X} - 68}{3.6/\sqrt{36}} \sim N(0, 1)$$

- 确定检验统计量为 Z

选择检验统计量

检验统计量定义

用于假设检验决策的统计量

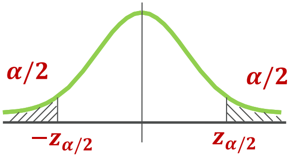
- 若原假设正确，则 $|\overline{x}-\mu|$ 不应太大，且

$$Z=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}=\frac{\overline{X}-68}{3.6/\sqrt{36}}\sim N(0,1)$$

- 确定检验统计量为 Z

确定拒绝域：显著性

- 若原假设正确，$|Z| = \left|\dfrac{\overline{X}-68}{0.6}\right|$ 取较大值是小概率事件。



- 确定一个常数 k 使得 $P\left(\left|\dfrac{\overline{X}-68}{0.6}\right| \geq k\right) = \alpha$
  - α 称为 显著性水平 ，常取 0.01, 0.05, 0.10 等。
  - 小概率事件的概率，由研究者确定。

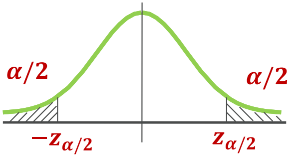
拒绝域和临界点

取 $\alpha = 0.05$，由标准正态分布分位点的定义得 $k = z_{\alpha/2} = z_{0.025} = 1.96$

- 当检验统计量取某个区域 C 中的值时，就拒绝原假设 H_0，则称区域 C 为拒绝域，拒绝域的边界点为临界点。
- $Z \geq z_{\alpha/2}$ 为 拒绝域 ，$\pm z_{0.025}$ 为 临界点 。

确定拒绝域：显著性

- 若原假设正确，$|Z| = \left|\frac{\overline{X}-68}{0.6}\right|$ 取较大值是小概率事件。
- 确定一个常数 k 使得 $P\left(\left|\frac{\overline{X}-68}{0.6}\right| \geq k\right) = \alpha$
  - α 称为 显著性水平 ，常取 0.01, 0.05, 0.10 等。
  - 小概率事件的概率，由研究者确定。



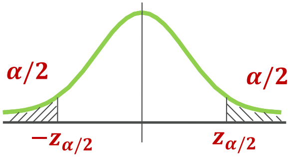
拒绝域和临界点

取 $\alpha = 0.05$，由标准正态分布分位点的定义得 $k = z_{\alpha/2} = z_{0.025} = 1.96$

- 当检验统计量取某个区域 C 中的值时，就拒绝原假设 H_0，则称区域 C 为拒绝域，拒绝域的边界点为临界点。
- $Z \geq z_{\alpha/2}$ 为 拒绝域 ，$\pm z_{0.025}$ 为 临界点 。

确定拒绝域：显著性

- 若原假设正确，$|Z| = \left|\frac{\overline{X}-68}{0.6}\right|$ 取较大值是小概率事件。
- 确定一个常数 k 使得 $P\left(\left|\frac{\overline{X}-68}{0.6}\right| \geq k\right) = \alpha$
  - α 称为 显著性水平 ，常取 0.01, 0.05, 0.10 等。
  - 小概率事件的概率，由研究者确定。

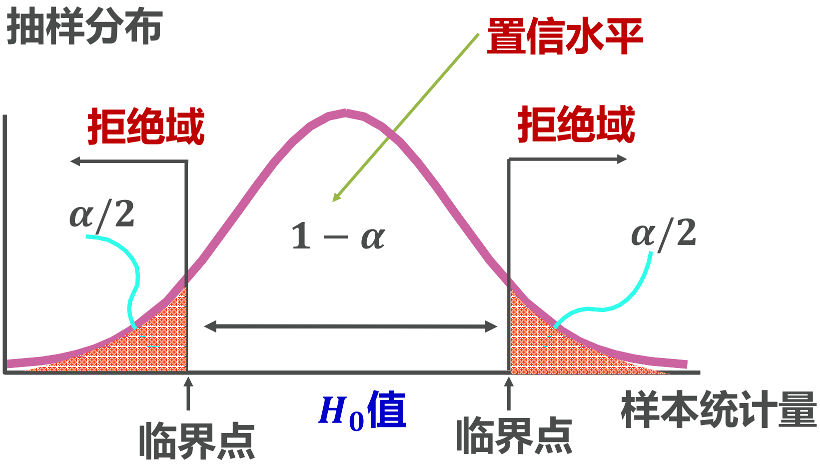


拒绝域和临界点

取 $\alpha = 0.05$, 由标准正态分布分位点的定义得 $k = z_{\alpha/2} = z_{0.025} = 1.96$

- 当检验统计量取某个区域 C 中的值时，就拒绝原假设 H_0，则称区域 C 为拒绝域，拒绝域的边界点为临界点。
- $Z \geq z_{\alpha/2}$ 为 拒绝域 ，$\pm z_{0.025}$ 为 临界点 。

定义总结



假设检验与区间估计的关系

- 要检验假设 $H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$，显著水平 α ，等价于计算 μ 的置信度为 $1-\alpha$ 的置信区间。
- 如果一个观测值 $\bar{x}$ 在显著性水平 α 下未能拒绝 H_0, 则意味着

$$\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| < z_{\alpha/2}$$

即 $\bar{x}$ 在置信度为 $1-\alpha$ 的置信区间中。

假设检验与区间估计的关系

- 要检验假设 $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$，显著水平 α ，等价于计算 μ 的置信度为 $1-\alpha$ 的置信区间。
- 如果一个观测值 $\overline{x}$ 在显著性水平 α 下未能拒绝 H_0, 则意味着

$$\left|\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right| < z_{\alpha/2}$$

即 $\overline{x}$ 在置信度为 $1-\alpha$ 的置信区间中。

计算下结论

1. 将检验统计量的值与 α 水平的临界值进行比较
2. 得出拒绝或不拒绝原假设的结论

引例

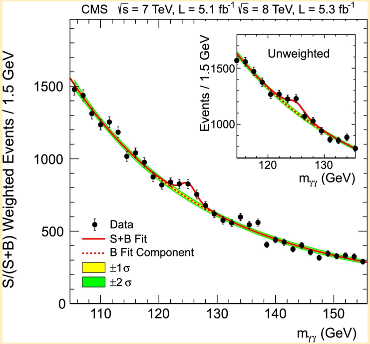
$\overline{x}=68.5$ ， $z=\left|\dfrac{68.5-68}{0.6}\right|=0.833<z_{0.025}=1.96$ 接受原假设，H_0：$\mu=68$。

- 结论：这批螺钉符合要求

四步曲

1. 根据实际问题所关心的内容，建立 H_0 与 H_1
2. 在 H_0 为真时，选择合适的统计量 Z
3. 由 H_1 确定拒绝域形式。给定显著性水平 α，确定拒绝域
4. 根据样本观测值计算，并作出相应的判断.

发现 Higgs： the god ~~damn~~ particle

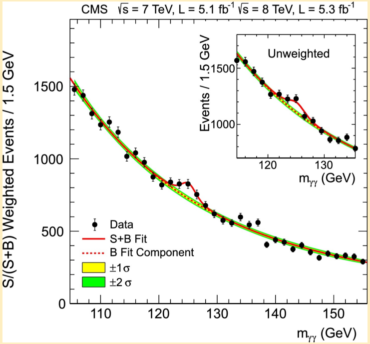


1. H_0: Higgs 的产生率为 0
   - H_1: Higgs 的产生率不为 0
2. 统计量 χ^2
3. 确定拒绝域 $\chi^2 > \chi^2_{0.00000029}(n)$ 5σ
4. 下结论：拒绝原假设

四步曲

1. 根据实际问题所关心的内容，建立 H_0 与 H_1
2. 在 H_0 为真时，选择合适的统计量 Z
3. 由 H_1 确定拒绝域形式。给定显著性水平 α，确定拒绝域
4. 根据样本观测值计算，并作出相应的判断.

发现 Higgs： the god ~~damn~~ particle

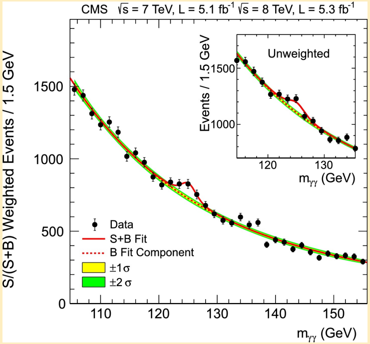


1. H_0: Higgs 的产生率为 0
   - H_1: Higgs 的产生率不为 0
2. 统计量 χ^2
3. 确定拒绝域 $\chi^2 > \chi^2_{0.00000029}(n)$ 5σ
4. 下结论：拒绝原假设

四步曲

1. 根据实际问题所关心的内容，建立 H_0 与 H_1
2. 在 H_0 为真时，选择合适的统计量 Z
3. 由 H_1 确定拒绝域形式。给定显著性水平 α ，确定拒绝域
4. 根据样本观测值计算，并作出相应的判断.

发现 Higgs： the god ~~damn~~ particle

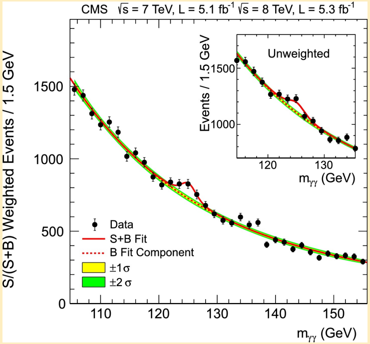


1. H_0: Higgs 的产生率为 0
   - H_1: Higgs 的产生率不为 0
2. 统计量 χ^2
3. 确定拒绝域 $\chi^2 > \chi^2_{0.00000029}(n)$ 5σ
4. 下结论：拒绝原假设

四步曲

1. 根据实际问题所关心的内容，建立 H_0 与 H_1
2. 在 H_0 为真时，选择合适的统计量 Z
3. 由 H_1 确定拒绝域形式。给定显著性水平 α ，确定拒绝域
4. 根据样本观测值计算，并作出相应的判断.

发现 Higgs： the god ~~damn~~ particle

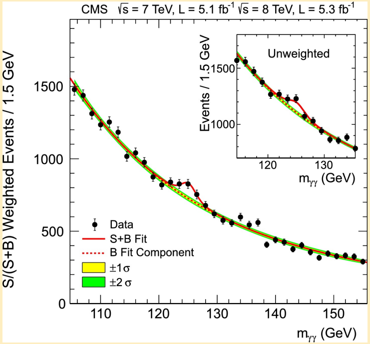


1. H_0: Higgs 的产生率为 0
   - H_1: Higgs 的产生率不为 0
2. 统计量 χ^2
3. 确定拒绝域 $\chi^2 > \chi^2_{0.00000029}(n)$ 5σ
4. 下结论：拒绝原假设

四步曲

1. 根据实际问题所关心的内容，建立 H_0 与 H_1
2. 在 H_0 为真时，选择合适的统计量 Z
3. 由 H_1 确定拒绝域形式。给定显著性水平 α，确定拒绝域
4. 根据样本观测值计算，并作出相应的判断.

发现 Higgs： the god ~~damn~~ particle

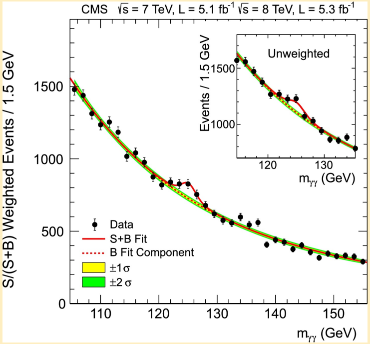


1. H_0: Higgs 的产生率为 0
   - H_1: Higgs 的产生率不为 0
2. 统计量 χ^2
3. 确定拒绝域 $\chi^2 > \chi^2_{0.00000029}(n)$ 5σ
4. 下结论：拒绝原假设

四步曲

1. 根据实际问题所关心的内容，建立 H_0 与 H_1
2. 在 H_0 为真时，选择合适的统计量 Z
3. 由 H_1 确定拒绝域形式。给定显著性水平 α，确定拒绝域
4. 根据样本观测值计算，并作出相应的判断.

发现 Higgs： the god ~~damn~~ particle

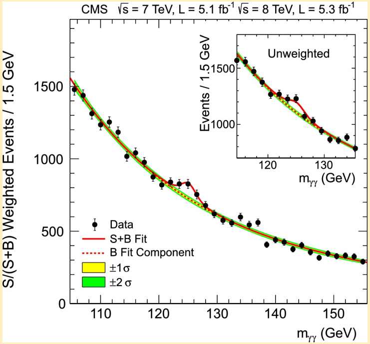


1. H_0: Higgs 的产生率为 0
   - H_1: Higgs 的产生率不为 0
2. 统计量 χ^2
3. 确定拒绝域 $\chi^2 > \chi^2_{0.00000029}(n)$ 5σ
4. 下结论：拒绝原假设

四步曲

1. 根据实际问题所关心的内容，建立 H_0 与 H_1
2. 在 H_0 为真时，选择合适的统计量 Z
3. 由 H_1 确定拒绝域形式。给定显著性水平 α ，确定拒绝域
4. 根据样本观测值计算，并作出相应的判断.

发现 Higgs： the god ~~damn~~ particle

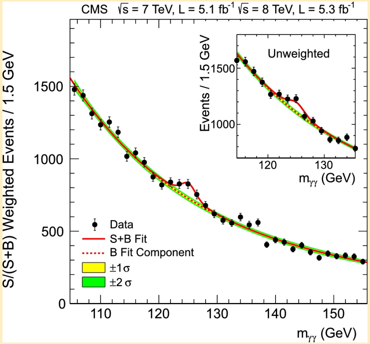


1. H_0: Higgs 的产生率为 0
   - H_1: Higgs 的产生率不为 0
2. 统计量 χ^2
3. 确定拒绝域 $\chi^2 > \chi^2_{0.00000029}(n)$ 5σ
4. 下结论：拒绝原假设

四步曲

1. 根据实际问题所关心的内容，建立 H_0 与 H_1
2. 在 H_0 为真时，选择合适的统计量 Z
3. 由 H_1 确定拒绝域形式。给定显著性水平 α ，确定拒绝域
4. 根据样本观测值计算，并作出相应的判断.

发现 Higgs： the god ~~damn~~ particle

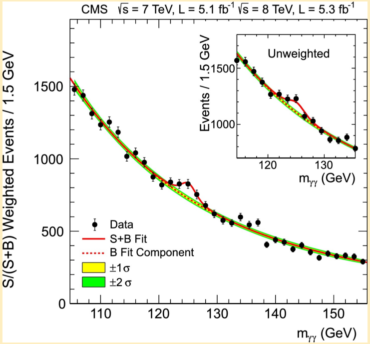


1. H_0: Higgs 的产生率为 0
   - H_1: Higgs 的产生率不为 0
2. 统计量 χ^2
3. 确定拒绝域 $\chi^2 > \chi^2_{0.00000029}(n)$ 5σ
4. 下结论：拒绝原假设

四步曲

1. 根据实际问题所关心的内容，建立 H_0 与 H_1
2. 在 H_0 为真时，选择合适的统计量 Z
3. 由 H_1 确定拒绝域形式。给定显著性水平 α，确定拒绝域
4. 根据样本观测值计算，并作出相应的判断.

发现 Higgs： the god ~~damn~~ particle



1. H_0: Higgs 的产生率为 0
   - H_1: Higgs 的产生率不为 0
2. 统计量 χ^2
3. 确定拒绝域 $\chi^2 > \chi^2_{0.00000029}(n)$ 5σ
4. 下结论：拒绝原假设

双边和单边

拒绝域的分类

假设	双边检验	左边检验	右边检验
H_0	$m = m_0$	$m \geq m_0$	$m \leq m_0$
H_1	$m \neq m_0$	$m < m_0$	$m > m_0$

双边检验例子

- 零件尺寸，既不能大也不能小

右边检验例子

- 采用新技术生产后，仪器寿命延长到 1500 小时

左边检验例子

- 改进生产工艺后，废品率降低到 2 % 以下

拒绝域的分类

假设	双边检验	左边检验	右边检验
H_0	$m = m_0$	$m \geq m_0$	$m \leq m_0$
H_1	$m \neq m_0$	$m < m_0$	$m > m_0$

双边检验例子

- 零件尺寸，既不能大也不能小

右边检验例子

- 采用新技术生产后，仪器寿命延长到 1500 小时

左边检验例子

- 改进生产工艺后，废品率降低到 2 % 以下

拒绝域的分类

假设	双边检验	左边检验	右边检验
H_0	$m = m_0$	$m \geq m_0$	$m \leq m_0$
H_1	$m \neq m_0$	$m < m_0$	$m > m_0$

双边检验例子

- 零件尺寸，既不能大也不能小

右边检验例子

- 采用新技术生产后，仪器寿命延长到 1500 小时

左边检验例子

- 改进生产工艺后，废品率降低到 2 % 以下

拒绝域的分类

假设	双边检验	左边检验	右边检验
H_0	$m = m_0$	$m \geq m_0$	$m \leq m_0$
H_1	$m \neq m_0$	$m < m_0$	$m > m_0$

双边检验例子

- 零件尺寸，既不能大也不能小

右边检验例子

- 采用新技术生产后，仪器寿命延长到 1500 小时

左边检验例子

- 改进生产工艺后，废品率降低到 2 % 以下

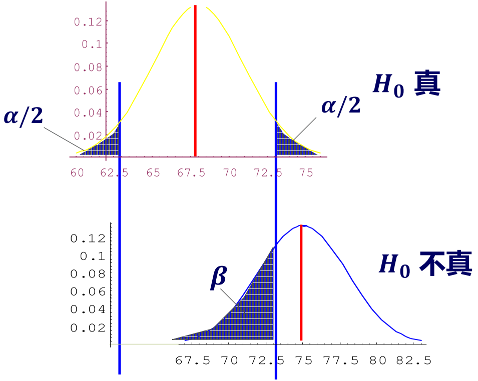
两类错误

假设检验的取舍

小概率事件在一次试验中很难发生，但很难发生不等于不发生，因而假设检验所作出的结论有可能是错误的。

❶ 第一类错误（弃真错误）
- 原假设为真时拒绝原假设
- 第一类错误的概率为显著性水平 α

❷ 第二类错误（取伪错误）
- 原假设为假时接受原假设
- 第二类错误的概率为 β



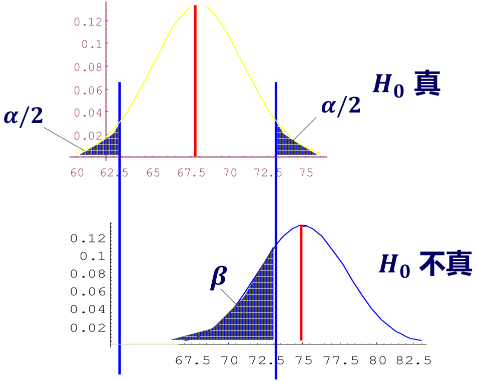
理想的检验方法应使犯两类错误的概率都很小，但在样本容量给定的情形下，不可能使两者都很小。降低一个，往往使另一个增大。

- 控制第一类错误的概率不超过 α, 然后, 若有必要, 通过增大样本容量的方法来减少 β .

假设检验的取舍

小概率事件在一次试验中很难发生，但很难发生不等于不发生，因而假设检验所作出的结论有可能是错误的。

1. 第一类错误（弃真错误）
   - 原假设为真时拒绝原假设
   - 第一类错误的概率为显著性水平 α
2. 第二类错误（取伪错误）
   - 原假设为假时接受原假设
   - 第二类错误的概率为 β



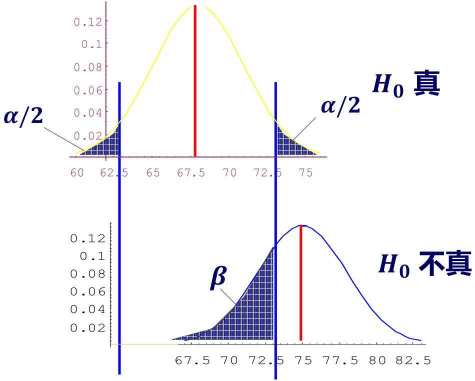
理想的检验方法应使犯两类错误的概率都很小，但在样本容量给定的情形下，不可能使两者都很小。降低一个，往往使另一个增大。

- 控制第一类错误的概率不超过 α, 然后, 若有必要, 通过增大样本容量的方法来减少 β .

假设检验的取舍

小概率事件在一次试验中很难发生，但很难发生不等于不发生，因而假设检验所作出的结论有可能是错误的。

1. 第一类错误（弃真错误）
   - 原假设为真时拒绝原假设
   - 第一类错误的概率为显著性水平 α
2. 第二类错误（取伪错误）
   - 原假设为假时接受原假设
   - 第二类错误的概率为 β



理想的检验方法应使犯两类错误的概率都很小，但在样本容量给定的情形下，不可能使两者都很小。降低一个，往往使另一个增大。

- 控制第一类错误的概率不超过 α，然后，若有必要，通过增大样本容量的方法来减少 β .

正态总体

σ^2 已知，关于 μ 的 Z 检验

正态总体 $N(\mu, \sigma^2)$ 当 σ^2 已知时，关于 $\mu = \mu_0$ 的检验问题：

1. $H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$
2. 选择统计量 $Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$，当 H_0 成立时，$Z \sim N(0, 1)$。
3. 对于给定的显著性水平 $\alpha (0 < \alpha < 1)$ 由标准正态分布分位数定义知 $P(|Z| \geq z_{\alpha/2}) = \alpha$。因此，检验的拒绝域为
$$C = \{|Z| \geq z_{\alpha/2}\}$$
4. 这种利用 Z 统计量来检验的方法称为 Z 检验法。

单边检验

- 右边检验
$$H_0 : \mu \leq \mu_0, H_1 : \mu > \mu_0, C = \{Z \geq z_\alpha\}$$
- 左边检验
$$H_0 : \mu \geq \mu_0, H_1 : \mu < \mu_0, C = \{Z \leq -z_\alpha\}$$

σ^2 已知，关于 μ 的 Z 检验

正态总体 $N(\mu,\sigma^2)$ 当 σ^2 已知时，关于 $\mu=\mu_0$ 的检验问题：

1. $H_0:\mu=\mu_0, H_1:\mu\neq\mu_0$
2. 选择统计量 $Z=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$，当 H_0 成立时，$Z\sim N(0,1)$。
3. 对于给定的显著性水平 $\alpha(0<\alpha<1)$ 由标准正态分布分位数定义知 $P(|Z|\geq z_{\alpha/2})=\alpha$。因此，检验的拒绝域为

$$C=\{|Z|\geq z_{\alpha/2}\}$$

4. 这种利用 Z 统计量来检验的方法称为 Z 检验法。

单边检验

- 右边检验

$$H_0:\mu\leq\mu_0, H_1:\mu>\mu_0, C=\{Z\geq z_\alpha\}$$

- 左边检验

$$H_0:\mu\geq\mu_0, H_1:\mu<\mu_0, C=\{Z\leq -z_\alpha\}$$

σ^2 已知，关于 μ 的 Z 检验

正态总体 $N(\mu, \sigma^2)$ 当 σ^2 已知时，关于 $\mu = \mu_0$ 的检验问题：

1. $H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$
2. 选择统计量 $Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$，当 H_0 成立时，$Z \sim N(0, 1)$。
3. 对于给定的显著性水平 $\alpha(0 < \alpha < 1)$ 由标准正态分布分位数定义知 $P(|Z| \geq z_{\alpha/2}) = \alpha$。因此，检验的拒绝域为

$$C = \{|Z| \geq z_{\alpha/2}\}$$

4. 这种利用 Z 统计量来检验的方法称为 Z 检验法。

单边检验

- 右边检验

$$H_0 : \mu \leq \mu_0, H_1 : \mu > \mu_0, C = \{Z \geq z_\alpha\}$$

- 左边检验

$$H_0 : \mu \geq \mu_0, H_1 : \mu < \mu_0, C = \{Z \leq -z_\alpha\}$$

σ^2 已知，关于 μ 的 Z 检验

正态总体 $N(\mu,\sigma^2)$ 当 σ^2 已知时，关于 $\mu=\mu_0$ 的检验问题：

1. $H_0:\mu=\mu_0, H_1:\mu\neq\mu_0$
2. 选择统计量 $Z=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$，当 H_0 成立时，$Z\sim N(0,1)$。
3. 对于给定的显著性水平 $\alpha(0<\alpha<1)$ 由标准正态分布分位数定义知 $P(|Z|\geq z_{\alpha/2})=\alpha$。因此，检验的拒绝域为
$$C=\{|Z|\geq z_{\alpha/2}\}$$
4. 这种利用 Z 统计量来检验的方法称为 Z 检验法。

单边检验

- 右边检验
$$H_0:\mu\leq\mu_0, H_1:\mu>\mu_0, C=\{Z\geq z_\alpha\}$$
- 左边检验
$$H_0:\mu\geq\mu_0, H_1:\mu<\mu_0, C=\{Z\leq -z_\alpha\}$$

σ^2 已知，关于 μ 的 Z 检验

正态总体 $N(\mu, \sigma^2)$ 当 σ^2 已知时，关于 $\mu = \mu_0$ 的检验问题：

1. $H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$
2. 选择统计量 $Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$，当 H_0 成立时，$Z \sim N(0, 1)$。
3. 对于给定的显著性水平 $\alpha(0 < \alpha < 1)$ 由标准正态分布分位数定义知 $P(|Z| \geq z_{\alpha/2}) = \alpha$。因此，检验的拒绝域为
$$C = \{|Z| \geq z_{\alpha/2}\}$$
4. 这种利用 Z 统计量来检验的方法称为 Z 检验法。

单边检验

- 右边检验
$$H_0 : \mu \leq \mu_0, H_1 : \mu > \mu_0, C = \{Z \geq z_\alpha\}$$
- 左边检验
$$H_0 : \mu \geq \mu_0, H_1 : \mu < \mu_0, C = \{Z \leq -z_\alpha\}$$

σ^2 已知，关于 μ 的 Z 检验

正态总体 $N(\mu,\sigma^2)$ 当 σ^2 已知时，关于 $\mu=\mu_0$ 的检验问题：

1. $H_0:\mu=\mu_0, H_1\colon\mu\neq\mu_0$
2. 选择统计量 $Z=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$，当 H_0 成立时，$Z\sim N(0,1)$。
3. 对于给定的显著性水平 $\alpha(0<\alpha<1)$ 由标准正态分布分位数定义知 $P(|Z|\geq z_{\alpha/2})=\alpha$。因此，检验的拒绝域为
$$C=\{|Z|\geq z_{\alpha/2}\}$$
4. 这种利用 Z 统计量来检验的方法称为 Z 检验法。

单边检验

- 右边检验
$$H_0\colon\mu\leq\mu_0, H_1\colon\mu>\mu_0, C=\{Z\geq z_\alpha\}$$
- 左边检验
$$H_0\colon\mu\geq\mu_0, H_1\colon\mu<\mu_0, C=\{Z\leq -z_\alpha\}$$

σ^2 为未知，关于 μ 的 t 检验

- 设总体 $X \sim N(\mu, \sigma^2)$，其中 μ, σ^2 未知。
- 设 $X_1, X_2, \cdots, X_n$ 为来自总体 X 的样本。
  1. 检验假设 $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$。显著水平为 α 。
  2. 采用 $T = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$ 作为检验统计量。
  3. 当 H_0 为真时，$T \sim t(n-1)$，由 t 分布分位点的定义知

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

     拒绝域为

     $$C = \{|T| \geq t_{\alpha/2}(n-1)\}$$
  4. 利用 t 统计量得出的检验法称为 t 检验法．正态总体的方差常为未知时，常用于检验总体均值。

σ^2 为未知，关于 μ 的 t 检验

- 设总体 $X \sim N(\mu, \sigma^2)$，其中 μ, σ^2 未知。
- 设 $X_1, X_2, \cdots, X_n$ 为来自总体 X 的样本。
  1. 检验假设 H_0:$\mu = \mu_0$,H_1:$\mu \neq \mu_0$。显著水平为 α 。
  2. 采用 $T = \frac{\overline{X}-\mu_0}{S/\sqrt{n}}$ 作为检验统计量。
  3. 当 H_0 为真时，$T \sim t(n-1)$，由 t 分布分位点的定义知
  $$P\left(\left|\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\right| \geq t_{\alpha/2}(n-1)\right) = \alpha$$
  拒绝域为
  $$C = \{|T| \geq t_{\alpha/2}(n-1)\}$$
  4. 利用 t 统计量得出的检验法称为 t 检验法. 正态总体的方差常为未知时，常用于检验总体均值。

σ^2 为未知，关于 μ 的 t 检验

- 设总体 $X \sim N(\mu, \sigma^2)$，其中 μ, σ^2 未知。
- 设 $X_1, X_2, \cdots, X_n$ 为来自总体 X 的样本。

1. 检验假设 $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$。显著水平为 α 。
2. 采用 $T = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$ 作为检验统计量。
3. 当 H_0 为真时，$T \sim t(n-1)$，由 t 分布分位点的定义知

$$P\left(\left|\frac{X - \mu_0}{S/\sqrt{n}}\right| \geq t_{\alpha/2}(n-1)\right) = \alpha$$

拒绝域为

$$C = \{|T| \geq t_{\alpha/2}(n-1)\}$$

4. 利用 t 统计量得出的检验法称为 t 检验法. 正态总体的方差常为未知时，常用于检验总体均值。

σ^2 为未知，关于 μ 的 t 检验

- 设总体 $X \sim N(\mu, \sigma^2)$，其中 μ, σ^2 未知。
- 设 $X_1, X_2, \cdots, X_n$ 为来自总体 X 的样本。
  1. 检验假设 $H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$。显著水平为 α 。
  2. 采用 $T = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$ 作为检验统计量。
  3. 当 H_0 为真时，$T \sim t(n-1)$，由 t 分布分位点的定义知

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

     拒绝域为

     $$C = \{|T| \geq t_{\alpha/2}(n-1)\}$$

  4. 利用 t 统计量得出的检验法称为 t 检验法. 正态总体的方差常为未知时，常用于检验总体均值。

μ 未知，关于 σ^2 的 χ^2 检验

设总体 $X \sim N(\mu, \sigma^2)$，其中 μ, σ^2 未知。$X_1, X_2, \cdots, X_n$ 为来自总体 X 的样本。

1. 检验假设 $H_0 \colon \sigma^2 = \sigma_0^2, H_1 \colon \sigma^2 \neq \sigma_0^2$。显著水平为 α。
2. 采用 $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$ 为检验统计量。
3. 当 H_0 为真时，由 χ^2 分布分位数的定义知

$$P(\chi^2 \leq \chi^2_{1-\alpha/2}(n-1)) = P(\chi^2 \geq \chi^2_{\alpha/2}(n-1)) = \alpha/2$$

拒绝域为

$$(-\infty, \chi^2_{1-\alpha/2}(n-1)] \cup [\chi^2_{\alpha/2}(n-1), +\infty)$$

μ 未知，关于 σ^2 的 χ^2 检验

设总体 $X \sim N(\mu, \sigma^2)$，其中 μ, σ^2 未知。$X_1, X_2, \cdots, X_n$ 为来自总体 X 的样本。

1. 检验假设 $H_0 : \sigma^2 = \sigma_0^2, H_1 : \sigma^2 \neq \sigma_0^2$。显著水平为 α。
2. 采用 $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$ 为检验统计量。
3. 当 H_0 为真时，由 χ^2 分布分位数的定义知

$$P(\chi^2 \leq \chi^2_{1-\alpha/2}(n-1)) = P(\chi^2 \geq \chi^2_{\alpha/2}(n-1)) = \alpha/2$$

拒绝域为

$$(-\infty, \chi^2_{1-\alpha/2}(n-1)] \cup [\chi^2_{\alpha/2}(n-1), +\infty)$$

μ 未知，关于 σ^2 的 χ^2 检验

设总体 $X \sim N(\mu, \sigma^2)$，其中 μ, σ^2 未知。$X_1, X_2, \cdots, X_n$ 为来自总体 X 的样本。

1. 检验假设 $H_0 : \sigma^2 = \sigma_0^2, H_1 : \sigma^2 \neq \sigma_0^2$。显著水平为 α。
2. 采用 $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$ 为检验统计量。
3. 当 H_0 为真时，由 χ^2 分布分位数的定义知

$$P(\chi^2 \leq \chi^2_{1-\alpha/2}(n-1)) = P(\chi^2 \geq \chi^2_{\alpha/2}(n-1)) = \alpha/2$$

拒绝域为

$$(-\infty, \chi^2_{1-\alpha/2}(n-1)] \cup [\chi^2_{\alpha/2}(n-1), +\infty)$$

小结

1. 单个总体均值 μ： Z 检验；t 检验
2. 单个正态总体方差：χ^2 检验

	原假设 H_0	检验统计量	备择假设 H_1	拒绝域
1	$\mu \le \mu_0$ $\mu \ge \mu_0$ $\mu = \mu_0$	$Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \ne \mu_0$	$z \ge z_\alpha$ $z \le -z_\alpha$ $\lvert z\rvert \ge z_{\alpha/2}$
2	$\mu \le \mu_0$ $\mu \ge \mu_0$ $\mu = \mu_0$ (σ^2未知)	$t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \ne \mu_0$	$t \ge t_\alpha(n-1)$ $t \le -t_\alpha(n-1)$ $\lvert t\rvert \ge t_{\alpha/2}(n-1)$
3	$\sigma^2 \le \sigma_0^2$ $\sigma^2 \ge \sigma_0^2$ $\sigma^2 = \sigma_0^2$ (μ未知)	$\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$	$\sigma^2 > \sigma_0^2$ $\sigma^2 < \sigma_0^2$ $\sigma^2 \ne \sigma_0^2$	$\chi^2 \ge \chi^2_\alpha(n-1)$ $\chi^2 \le \chi^2_{1-\alpha}(n-1)$ $\chi^2 \ge \chi^2_{\alpha/2}(n-1)$或 $\chi^2 \le \chi^2_{1-\alpha/2}(n-1)$

χ^2 检验

分布拟合 χ^2 检验

观察到了实验数据，可以用泊松分布描述吗？从样本来检验关于分布的假设。

1. 设总体 X 的分布未知， $x_1, x_2, \cdots, x_n$ 是来自 X 的样本值。检验假设

   H_0 总体 X 的分布函数为 $F(x)$
   H_1 总体 X 的分布函数不是 $F(x)$

2. 构造统计量。将在 H_0 下 X 可能取值的全体 Ω 分成互不相交的子集 $A_1, A_2, \cdots, A_k$ 。以 f_i 记样本观测值落在 A_i 的频数，$p_i = P(A_i)$。
   - 皮尔逊 Pearson 证明，选取

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

   近似服从 $\chi^2(k-1)$ 分布

3. 设定拒绝域。当 H_0 为真时，χ^2 不应太大，若过大就拒绝 H_0。拒绝域的形式为

   $$\chi^2 \geq \chi^2_\alpha(k-1)$$

分布拟合 χ^2 检验

观察到了实验数据，可以用泊松分布描述吗？从样本来检验关于分布的假设。

1. 设总体 X 的分布未知， $x_1, x_2, \cdots, x_n$ 是来自 X 的样本值。检验假设

   H_0 总体 X 的分布函数为 $F(x)$

   H_1 总体 X 的分布函数不是 $F(x)$

2. 构造统计量。将在 H_0 下 X 可能取值的全体 Ω 分成互不相交的子集 $A_1, A_2, \cdots, A_k$ 。以 f_i 记样本观测值落在 A_i 的频数，$p_i = P(A_i)$。
   - 皮尔逊 Pearson 证明，选取

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

   近似服从 $\chi^2(k-1)$ 分布

3. 设定拒绝域。当 H_0 为真时，χ^2 不应太大，若过大就拒绝 H_0。拒绝域的形式为

   $$\chi^2 \geq \chi^2_\alpha(k-1)$$

分布拟合 χ^2 检验

观察到了实验数据，可以用泊松分布描述吗？从样本来检验关于分布的假设。

1. 设总体 X 的分布未知， $x_1, x_2, \cdots, x_n$ 是来自 X 的样本值。检验假设

   H_0 总体 X 的分布函数为 $F(x)$
   H_1 总体 X 的分布函数不是 $F(x)$

2. 构造统计量。将在 H_0 下 X 可能取值的全体 Ω 分成互不相交的子集 $A_1, A_2, \cdots, A_k$ 。以 f_i 记样本观测值落在 A_i 的频数，$p_i = P(A_i)$。
   - 皮尔逊 Pearson 证明，选取

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

   近似服从 $\chi^2(k-1)$ 分布

3. 设定拒绝域。当 H_0 为真时，χ^2 不应太大，若过大就拒绝 H_0。拒绝域的形式为

   $$\chi^2 \geq \chi^2_\alpha(k-1)$$

χ^2 检验条件

$n \geq 50, np_i \geq 5$

分布族的 χ^2 拟合检验

如果遇到的原假设中有 r 个未知参数 $\vec{\theta}$ ，则先在 H_0 的假设下，对参数进行估计得 $\hat{\vec{\theta}}$，从而得到 $\hat{p}_i = P(A_i; \hat{\vec{\theta}})$。取

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

作为检验统计量，近似服从 $\chi^2(k-r-1)$ 。

参数估计 使用估计量推断未知参数

假设检验 对参数有所了解 $\rightarrow$ 但有怀疑猜测需要证实之时

例：天然放射性的泊松分布检验

例

每隔一天观察一次铀样品被计数的 α 个数 X ，共观察了 100 天，结果如下：

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

1. 检验 $H_0: X$ 服从泊松分布
2. 检验统计量：
   - 先由最大似然估计得 $\hat{\lambda} = \overline{x} = 4.2$。
   - 对天数进行分组，保证 $\hat{p}_i n \geq 5$

i	0-1	2	3	4	5	6	7	≥ 8
f_i	6	16	17	26	11	9	9	6

例：天然放射性的泊松分布检验

例

每隔一天观察一次铀样品被计数的 α 个数 X ，共观察了 100 天，结果如下：

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

1. 检验 H_0 : X 服从泊松分布
2. 检验统计量：
    - 先由最大似然估计得 $\hat{\lambda} = \overline{x} = 4.2$。
    - 对天数进行分组，保证 $\hat{p}_i n \geq 5$

i	0-1	2	3	4	5	6	7	≥ 8
f_i	6	16	17	26	11	9	9	6

天然放射性的泊松分布检验计算

```
observed <- c(6, 16, 17, 26, 11, 9, 9, 6)
expected <- c(ppois(1, 4.2), dpois(2, 4.2), dpois(3, 4.2), dpois(4, 4.2),
             dpois(5, 4.2), dpois(6, 4.2), dpois(7, 4.2), 1-ppois(7, 4.2))
chisq.statistic <- chisq.test(observed, p=expected)$statistic
chisq.statistic
```

```
6.28260565272422
```

- 自由度为 6，0.05 显著度下的临界值

```
qchisq(1-0.05, df=6)
```

```
12.591587243744
```

- 6.28 不在拒绝域内，接受原假设。X 服从泊松分布。

p 值检验

定义

p 值

假设检验问题的 p 值（probability value，p-value）是由检验统计量的样本观察值得出的原假设可被拒绝的**最小显著性水平**。

任一检验问题的 p 值可以根据检验统计量的样本观察值的以及检验统计量在 H_0 下一个特定的参数值（一般是 H_0 与 H_1 所规定的参数的分界点）对应的分布求出。

使用

对于任意指定的显著性水平 α，

- 若 p值 $\leq \alpha$，则在显著性水平 α 下拒绝 H_0；
- 若 p值 $> \alpha$，则在显著性水平 α 下接受 H_0；

利用 p 值确定是否拒绝 H_0 的方法，称为 **p 值法**。

定义

p 值

假设检验问题的 p 值（probability value，p-value）是由检验统计量的样本观察值得出的原假设可被拒绝的最小显著性水平。

任一检验问题的 p 值可以根据检验统计量的样本观察值的以及检验统计量在 H_0 下一个特定的参数值（一般是 H_0 与 H_1 所规定的参数的分界点）对应的分布求出。

使用

对于任意指定的显著性水平 α，

- 若 p值 $\leq \alpha$，则在显著性水平 α 下拒绝 H_0；
- 若 p值 $> \alpha$，则在显著性水平 α 下接受 H_0；

利用 p 值确定是否拒绝 H_0 的方法，称为 p 值法。

定义

p 值

假设检验问题的 p 值（probability value，p-value）是由检验统计量的样本观察值得出的原假设可被拒绝的 最小显著性水平 。

任一检验问题的 p 值可以根据检验统计量的样本观察值的以及检验统计量在 H_0 下一个特定的参数值（一般是 H_0 与 H_1 所规定的参数的分界点）对应的分布求出。

使用

对于任意指定的显著性水平 α，

- 若 p值 $\leq \alpha$，则在显著性水平 α 下拒绝 H_0 ；
- 若 p值 $> \alpha$，则在显著性水平 α 下接受 H_0 ；

利用 p 值确定是否拒绝 H_0 的方法，称为 p 值法 。

例：轴样品的 α 计数

```
1 - pchisq(chisq.statistic, df=6)
```

0.392288394887097

- p 值为 0.39。大于 0.05 的显著性水平，接受原假设。

```
x <- seq(0,30,length.out=100)
par(mfrow = c(1, 2))
plot(x, dchisq(x, df=6), type='l', ylab="概率密度")
sel.x <- x[chisq.statistic < x]
polygon(c(sel.x[1], sel.x), c(0, dchisq(sel.x, df=6)), col="gray")
plot(x, pchisq(x, df=6), type='l', ylab="累积分布")
```