GB

中国国家标准汇编

2008年 修订-20

中国国家标准汇编

2008年修订-20

中国标准出版社　编

中国标准出版社
北京

图书在版编目（CIP）数据

中国国家标准汇编：2008 年修订．20/中国标准出版社编．—北京：中国标准出版社，2009

ISBN 978-7-5066-5363-3

Ⅰ．中…　Ⅱ．中…　Ⅲ．国家标准-汇编-中国-2008
Ⅳ．T-652.1

中国版本图书馆 CIP 数据核字（2009）第 102486 号

中 国 标 准 出 版 社 出 版 发 行
北京复兴门外三里河北街 16 号
邮政编码：100045
网址 www.spc.net.cn
电话：68523946　68517548
中国标准出版社秦皇岛印刷厂印刷
各地新华书店经销
*
开本 880×1230　1/16　印张 38.75　字数 1 169 千字
2009 年 7 月第一版　2009 年 7 月第一次印刷
*
定价 200.00 元

ISBN 978-7-5066-5363-3

出 版 说 明

1.《中国国家标准汇编》是一部大型综合性国家标准全集。自1983年起，按国家标准顺序号以精装本、平装本两种装帧形式陆续分册汇编出版。它在一定程度上反映了我国建国以来标准化事业发展的基本情况和主要成就，是各级标准化管理机构，工矿企事业单位，农林牧副渔系统，科研、设计、教学等部门必不可少的工具书。

2.《中国国家标准汇编》收入我国每年正式发布的全部国家标准。分为“制定”卷和“修订”卷两种编辑版本。

“制定”卷收入上年度我国发布的、新制定的国家标准，顺延前年度标准编号分成若干分册，封面和书脊上注明“20××年制定”字样及分册号，分册号一直连续。各分册中的标准是按照标准编号顺序连续排列的，如有标准顺序号缺号的，除特殊情况注明外，暂为空号。

“修订”卷收入上年度我国发布的、被修订的国家标准，视篇幅分设若干分册，但与“制定”卷分册号无关联，仅在封面和书脊上注明“20××年修订-1，-2，-3，……”字样。“修订”卷各分册中的标准，仍按标准编号顺序排列(但不连续)；如有遗漏的，均在当年最后一分册中补齐。需提请读者注意的是，个别非顺延前年度标准编号的新制定的国家标准没有收入在“制定”卷中，而是收入在“修订”卷中。

读者配套购买《中国国家标准汇编》“制定”卷和“修订”卷则可收齐上一年度我国制定和修订的全部国家标准。

3. 由于读者需求的变化，自1996年起，《中国国家标准汇编》仅出版精装本。

4. 2008年制修订国家标准共5 946项，本分册为“2008年修订-20”，收入新制修订的国家标准23项。

中国标准出版社

2009年5月

目　　录

ICS 03.120.30
A 41

中华人民共和国国家标准

GB/T 4089—2008
代替 GB/T 4089—1983,GB/T 4090—1983

数据的统计处理和解释 泊松分布参数的估计和检验

Statistical interpretation of data—Estimation and hypothesis test of parameter in poisson distribution

2008-07-16 发布 2009-01-01 实施

中华人民共和国国家质量监督检验检疫总局
中国国家标准化管理委员会 发布

前　言

本标准是在 GB/T 4089—1983《数据的统计处理和解释　泊松分布参数的估计》和 GB/T 4090—1983《数据的统计处理和解释　泊松分布参数的检验》的基础上整合而成，本标准代替 GB/T 4089—1983 和 GB/T 4090—1983。本标准与 GB/T 4089—1983 和 GB/T 4090—1983 相比较，技术内容的变化主要包括：

——增加了术语、符号和定义；

——在附录 C 中增加了 P 值检验。

本标准的附录 A 为规范性附录，附录 B 和附录 C 均为资料性附录。

本标准由中国标准化研究院提出。

本标准由全国统计方法应用标准化技术委员会归口。

本标准起草单位：中国标准化研究院、广州市产品质量监督检验所、北京大学、无锡市产品质量监督检验所、福州春伦茶业有限公司。

本标准主要起草人：于振凡、丁文兴、邓穗兴、房祥忠、陈华英、陈玉忠、傅天龙。

本标准所代替标准的历次版本发布情况为：

——GB/T 4089—1983；

——GB/T 4090—1983。

引 言

从事科学研究、工农业制造以及管理工作都离不开数据，而对这些数据的整理、分析和解释都离不开统计方法。统计学是研究数字资料的整理、分析和正确解释的一门学科。人们各自从不同的来源取得各种数字资料，这些数字资料通常都是杂乱无章的，必须经过整理和简缩才能利用，使用完善的统计方法就可使数据整理、排列的有条有理，用图形或少量的几个重要参数，就可把一大堆数据的特征表达出来，这样既可避免不正确的解释，又可将获得满意数据的成本降到最低限度，提高了经济效益。

国家标准《数据的统计处理和解释》包含以下各项：

——统计容忍区间的确定(GB/T 3359)

——均值的估计和置信区间(GB/T 3360)

——在成对观测值情形下两个均值的比较(GB/T 3361)

——二项分布参数的估计与检验(GB/T 4088)

——泊松分布参数的估计和检验(GB/T 4089)

——正态性检验(GB/T 4882)

——正态样本离群值的判断和处理(GB/T 4883)

——正态分布均值和方差的估计与检验(GB/T 4889)

——正态分布均值和方差检验的功效(GB/T 4890)

——Ⅰ型极值分布样本离群值的判断和处理(GB/T 6380)

——伽玛分布（皮尔逊Ⅲ型分布）的参数估计(GB/T 8055)

——指数分布样本离群值的判断和处理(GB/T 8056)

《数据的统计处理和解释　泊松分布参数的估计和检验》尚无相应的国际标准。

数据的统计处理和解释 泊松分布参数的估计和检验

1 范围

本标准基于泊松分布总体，分布的概率质量函数为：

$$P\{X=x/\lambda\}=\frac{\lambda^x}{x!}e^{-\lambda} \quad x=0,1,2\cdots$$

式中 $\lambda>0$ 为分布参数。本标准依据来自总体的独立随机样本 $x_1,x_2,\cdots,x_n$，规定了参数 λ 的点估计、区间估计和检验的方法。当有充分理由确信总体服从泊松分布时，可采用本标准。

2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。凡是注日期的引用文件，其随后所有的修改单(不包括勘误的内容)或修订版均不适用于本标准，然而，鼓励根据本标准达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件，其最新版本适用于本标准。

GB/T 4086.1 统计分布数值表 正态分布

GB/T 4086.2 统计分布数值表 χ^2 分布

ISO 3534-1:2006 统计学词汇及符号 第1部分：一般统计术语与用于概率的术语

ISO 3534-2:2006 统计学词汇及符号 第2部分：应用统计

3 术语、定义和符号

ISO 3534-1:2006 和 ISO 3534-2:2006 确定的以及下列术语、定义和符号适用于本标准。为便于参考，某些术语、符号直接引自上述标准。

3.1 术语和定义

ISO 3534-1:2006 和 ISO 3534-2:2006 确定的术语和定义适用于本标准。

3.2 符号

n	样本量
λ	泊松分布的参数
λ_L	区间估计中 λ 的置信下限
λ_U	区间估计中 λ 的置信上限
α	显著性水平
β	第二类错误概率的指定值
α'	<显著性检验>第一类错误概率，原假设为真时，拒绝原假设的概率
β'	<显著性检验>第二类错误概率，当原假设错误时，没有拒绝原假设的概率
H_0	<显著性检验>原假设
H_1	<显著性检验>备择假设
$P(A)$	事件 A 的概率
T	样本观测值 $x_1,x_2,\cdots,x_n$ 的总和，$T=\sum_{i=1}^{n}x_i$
$1-\alpha$	<区间估计>置信水平

4 参数估计

4.1 点估计

λ 的估计量记为 $\hat{\lambda}$。

$$\hat{\lambda} = \frac{T}{n} = \overline{x} \qquad \cdots\cdots(1)$$

4.2 区间估计

4.2.1 置信区间的三种形式

λ 的置信区间常用的有三种形式：

a) 双侧置信区间(λ_L,λ_U)，这里 $0<\lambda_L<\lambda_U<+\infty$。

b) 仅有置信下限 λ_L 的单侧置信区间(λ_L,$+\infty$)，这里 $\lambda_L>0$。

c) 仅有置信上限 λ_U 的单侧置信区间(0,λ_U)，这里 $\lambda_U>0$。

选用何种类型的置信区间，应根据所研究问题的需要而定。

λ_L 和 λ_U 是观测值的两个函数，本标准采用的置信区间所满足的条件为：所求得的置信区间包含真正的 λ 值这个事件发生的频率约等于置信水平 $1-\alpha$。具体地说，

双侧区间应满足：$P\{\lambda_L<\lambda<\lambda_U\}=1-\alpha$

单侧置信区间应满足：

$$P\{\lambda_L<\lambda\} = 1-\alpha$$
$$P\{\lambda<\lambda_U\} = 1-\alpha$$

根据不同的要求，$1-\alpha$ 的数值通常从 0.90，0.95，0.99 中选取。

由于泊松分布的离散性，不一定都能求得使上述等式成立的 λ_L 和(或)λ_U，这时取 λ_L，λ_U 满足：

$P\{\lambda_L<\lambda<\lambda_U\}\geqslant 1-\alpha$ （双侧情形）

或 $P\{\lambda>\lambda_L\}\geqslant 1-\alpha$ （单侧下限）

或 $P\{\lambda<\lambda_U\}\geqslant 1-\alpha$ （单侧上限）

4.2.2 置信限的求法

4.2.2.1 单侧置信下限

$$\lambda_L = \frac{1}{2n}\chi^2_{\alpha}(2T) \qquad \cdots\cdots(2)$$

式中 $\chi^2_{\alpha}(2T)$表示自由度为 $2T$ 的 χ^2 分布的 α 分位数，其值可由 χ^2 分位数表查得(见 GB/T 4086.2)。

示例

某次试验中，$n=10$，$T=\sum_{i=1}^{10} x_i = 14$，取 $1-\alpha=0.95$。

从 χ^2 分位数表中查得，$\chi^2_{\alpha}(2T)=\chi^2_{0.05}(28)=16.927\,9$，得单侧置信下限：

$$\lambda_L = \frac{16.927\,9}{20} = 0.846$$

所以单侧置信区间为(0.846，$+\infty$)。

4.2.2.2 单侧置信上限

$$\lambda_U = \frac{1}{2n}\chi^2_{1-\alpha}(2T+2) \qquad \cdots\cdots(3)$$

式中 $\chi^2_{1-\alpha}(2T+2)$表示自由度为 $2T+2$ 的 χ^2 分布的 $1-\alpha$ 分位数。

示例

某次试验中，$n=10$，$T=\sum_{i=1}^{10} x_i = 14$，取 $1-\alpha=0.95$。

从 χ^2 分位数表中查得，$\chi^2_{1-\alpha}(2T+2)=\chi^2_{0.95}(30)=43.773\,0$，则单侧置信上限：

$$\lambda_U = \frac{43.773\ 0}{20} = 2.189$$

所以单侧置信区间为(0,2.189)。

4.2.2.3 双侧置信区间的置信限

置信下限：$\lambda_L = \frac{1}{2n}\chi^2_{\alpha/2}(2T)$ …………（4）

置信上限：$\lambda_U = \frac{1}{2n}\chi^2_{1-\alpha/2}(2T+2)$ …………（5）

式中 $\chi^2_{\alpha/2}(2T)$表示自由度为 $2T$ 的 χ^2 分布的 $\alpha/2$ 分位数。$\chi^2_{1-\alpha/2}(2T+2)$表示自由度为 $2T+2$ 的 χ^2 分布的 $1-\alpha/2$ 分位数。

示例

某次试验中，$n=10$，$T=\sum_{i=1}^{10} x_i = 14$，取 $1-\alpha=0.95$。

从 χ^2 分位数表中查得，$\chi^2_{\alpha/2}(2T)=\chi^2_{0.025}(28)=15.307\ 9$

双侧置信区间的下限：$\lambda_L=\frac{15.307\ 9}{20}=0.765$

从 χ^2 分位数表中查得，$\chi^2_{1-\alpha/2}(2T+2)=\chi^2_{0.975}(30)=46.979\ 2$

双侧置信区间的上限：$\lambda_U=\frac{46.979\ 2}{20}=2.349$

所以双侧置信区间为(0.765,2.349)。

4.2.3 置信区间的近似求法

当 $2T>30$ 时，可使用下面给出的近似求法。

4.2.3.1 单侧置信下限的近似求法

单侧置信下限的近似计算公式为：

$$\lambda_L = \frac{1}{n}\{c + (\sqrt{T+2c-0.5} - \frac{u_{1-\alpha}}{2})^2\} \quad \cdots\cdots(6)$$

式中 $c=\frac{u^2_{1-\alpha}+2}{36}$，$u_{1-\alpha}$为标准正态分布的 $1-\alpha$ 分位数。

示例

在某次试验中，$n=15$，$T=\sum_{i=1}^{15} x_i = 18$，取 $1-\alpha=0.95$。

查正态分位数表(见 GB/T 4086.1)：

$$u_{1-\alpha} = u_{0.95} = 1.644\ 85$$

$$c = \frac{u^2_{1-\alpha}+2}{36} = 0.130\ 71$$

$$\frac{u_{1-\alpha}}{2} = 0.822\ 43$$

$$\sqrt{T+2c-0.5} = 4.214\ 43$$

$$\begin{aligned}\lambda_L &= \frac{1}{n}\{c + (\sqrt{T+2c-0.5} - \frac{u_{1-\alpha}}{2})^2\} \\ &= \frac{1}{15}\{0.130\ 71 + (4.214\ 43 - 0.822\ 43)^2\} \\ &= 0.776\end{aligned}$$

置信水平为 0.95 的单侧置信区间为(0.776,$+\infty$)。

4.2.3.2 单侧置信上限的近似求法

单侧置信上限的近似公式为：

$$\lambda_U = \frac{1}{n}\{c + (\sqrt{T+2c+0.5} + \frac{u_{1-\alpha}}{2})^2\} \quad \cdots\cdots(7)$$

式中 c，$u_{1-\alpha}$同 4.2.3.1。

示例

在某次试验中，$n=15$，$T=\sum_{i=1}^{15} x_i = 18$，取 $1-\alpha=0.95$。

查正态分位数表(见 GB/T 4086.1)：

$$u_{0.95} = 1.644\ 85$$

计算：

$$c = \frac{u_{1-\alpha}^2 + 2}{36} = \frac{(1.644\ 85)^2 + 2}{36} = 0.130\ 71$$

$$\frac{u_{1-\alpha}}{2} = \frac{1.644\ 85}{2} = 0.822\ 43$$

$$\sqrt{T+2c+0.5} = 4.331\ 45$$

$$\begin{aligned}\lambda_U &= \frac{1}{n}\{c + (\sqrt{T+2c+0.5} + \frac{u_{1-\alpha}}{2})^2\} \\ &= \frac{1}{15}\{0.130\ 71 + (4.331\ 45 + 0.822\ 43)^2\} \\ &= 1.780\end{aligned}$$

置信水平为 0.95 的单侧置信区间为(0，1.780)。

4.2.3.3 双侧置信区间置信限的近似求法

置信下限的近似计算公式为：

$$\lambda_L = \frac{1}{n}\{c + (\sqrt{T+2c-0.5} - \frac{u_{1-\alpha/2}}{2})^2\} \quad \cdots\cdots(8)$$

置信上限的近似计算公式为：

$$\lambda_U = \frac{1}{n}\{c + (\sqrt{T+2c+0.5} + \frac{u_{1-\alpha/2}}{2})^2\} \quad \cdots\cdots(9)$$

式中 $c=\frac{u_{1-\alpha/2}^2+2}{36}$，$u_{1-\alpha/2}$为标准正态分布的 $1-\alpha/2$ 分位数。

示例

在某次试验中，$n=15$，$T=\sum_{i=1}^{15} x_i = 18$，取 $1-\alpha=0.95$

查正态分位数表(见 GB/T 4086.1)：

$$u_{1-\alpha/2} = u_{0.975} = 1.959\ 96$$

$$c = \frac{u_{1-\alpha/2}^2 + 2}{36} = \frac{(1.959\ 96)^2 + 2}{36} = 0.162\ 26$$

$$\frac{u_{1-\alpha/2}}{2} = 0.879\ 98$$

$$\sqrt{T+2c-0.5} = 4.221\ 91$$

$$\sqrt{T+2c+0.5} = 4.338\ 72$$

$$\begin{aligned}\lambda_L &= \frac{1}{n}\{c + (\sqrt{T+2c-0.5} - \frac{u_{1-\alpha/2}}{2})^2\} \\ &= \frac{1}{15}\{0.162\ 26 + (4.221\ 91 - 0.879\ 98)^2\} \\ &= 0.755\end{aligned}$$

$$\lambda_U = \frac{1}{n}\{c+(\sqrt{T+2c+0.5}+\frac{u_{1-\alpha/2}}{2})^2\}$$
$$= \frac{1}{15}\{0.162\ 26+(4.338\ 72+0.879\ 98)^2\}$$
$$= 1.826$$

置信水平为 0.95 的双侧置信区间为(0.755,1.826)。

5 泊松分布参数的检验

5.1 原假设与备择假设

用 H_0 表示原假设,H_1 表示备择假设。本标准处理常见的三种情况:

检验分为以下三种情况:

$H_0:\lambda=\lambda_0$, $H_1:\lambda\neq\lambda_0$ (双侧检验)

$H_0:\lambda\leqslant\lambda_0$, $H_1:\lambda>\lambda_0$ (单侧检验)

$H_0:\lambda\geqslant\lambda_0$, $H_1:\lambda<\lambda_0$ (单侧检验)

选用何种类型的检验,应根据所研究问题的需要而定。本标准正文所列出的是常用的经典的方法,本标准附录 C 也提供了等价的利用置信区间的方法和计算 P 值的方法。

5.2 双侧检验 $H_0:\lambda=\lambda_0$ $H_1:\lambda\neq\lambda_0$

5.2.1 实施步骤

a) 由 λ_0,样本量 n 及给定的显著性水平 α,确定拒绝域的临界值 c_1 和 c_2。(c_1 和 c_2 的确定见 5.2.2);

b) 计算 $T=\sum\limits_{i=1}^{n}x_i$ 的值;

c) 当 $T\leqslant c_1$ 或 $T\geqslant c_2$ 时,拒绝 H_0;当 $c_1<T<c_2$ 时,不拒绝 H_0。

5.2.2 拒绝域临界值 c_1 和 c_2 的确定

c_1 是满足下式的最大整数

$$P\{T\leqslant c_1 \mid n\lambda_0\}=\sum_{k=0}^{c_1}\frac{(n\lambda_0)^k}{k!}e^{-n\lambda_0}\leqslant\frac{\alpha}{2}$$

c_2 是满足下式的最小整数

$$P\{T\geqslant c_2 \mid n\lambda_0\}=\sum_{k=c_2}^{\infty}\frac{(n\lambda_0)^k}{k!}e^{-n\lambda_0}\leqslant\frac{\alpha}{2}$$

相应的第一类错误的概率和第二类错误的概率见附录 A。

c_1 和 c_2 也可用 χ^2 分布表法(见 5.2.2.1)或正态近似法(见 5.2.2.2)确定。

对某些 λ_0、n 和 α,c_1 和(或)c_2 的值可能不存在。如当 $n\lambda_0$ 比较小,使得 $P\{T=0|n\lambda_0\}=e^{-n\lambda_0}>\frac{\alpha}{2}$ 时,则 c_1 就不存在。此时可通过协商适当加大 α 值。

5.2.2.1 χ^2 分布表法

c_1 是满足下式的最大整数:

$$2n\lambda_0\geqslant\chi^2_{1-\alpha/2}(2c_1+2) \quad\cdots\cdots(10)$$

式中 $\chi^2_{1-\alpha/2}(2c_1+2)$ 是自由度为 $2c_1+2$ 的 χ^2 分布的 $1-\alpha/2$ 分位数。

c_2 是满足下式的最小整数:

$$2n\lambda_0\leqslant\chi^2_{\alpha/2}(2c_2) \quad\cdots\cdots(11)$$

式中 $\chi^2_{\alpha/2}(2c_2)$ 是自由度为 $2c_2$ 的 χ^2 分布的 $\alpha/2$ 分位数。式中的分位数可以查 χ^2 分布分位数表。

示例

放射性物质在某一定长的时间间隔内所放射的粒子数 X 服从泊松分布。现观测某种放射性物质放出的 α 粒子数,一共作了 15 次观测,每次观测的时间为 90 s,观测结果列于表 1:

表 1

放射粒子数	观测到的频数
0	4
1	7
2	2
3	1
4	1
大于 4	0
总计	15

现利用这 15 次观测的结果，检验泊松分布的参数是否为 $\lambda=0.6$。检验的显著性水平取 $\alpha=0.10$。按照 5.2.1，检验步骤如表 2。

表 2

确定 H_0 和 H_1	$H_0:\lambda=0.6$， $H_1:\lambda\neq0.6$
a. 由 λ_0、n 和 α 确定 c_1 和 c_2。	由于 $\lambda_0=0.6$ $n=15$ 和 $\alpha=0.10$，得 $2n\lambda_0=18$， $\frac{\alpha}{2}=0.05, 1-\frac{\alpha}{2}=0.95$。 查 χ^2 分布分位数表得 $\chi^2_{0.95}(8)=15.597$ $\chi^2_{0.95}(10)=18.307$， 所以 $2c_1+2=8, c_1=3$ 又查 χ^2 分布分位数表得 $\chi^2_{0.05}(28)=16.928$， $\chi^2_{0.05}(30)=18.493$， 所以 $2c_2=30, c_2=15$
b. 计算 $T=\sum_{i=1}^{n}x_i$	$T=0\times4+1\times7+2\times2+3\times1+4\times1=18$
c. 判断	由于 $T=18>c_2=15$ 所以不拒绝原假设 $H_0:\lambda=0.6$

5.2.2.2 正态近似法

当 $n\lambda_0$ 较大时，所确定的 c_1 和 c_2 也将比较大。当用 χ^2 分布表法确定 c_1 和 c_2 有困难时，可用下述正态近似法，实际使用时优先采用泊松分布法或 χ^2 分布表法。

c_1 是满足下式的最大整数：

$$c_1+1\leqslant\left[\sqrt{n\lambda_0+\frac{4-u_{1-\alpha/2}^2}{12}}-\frac{u_{1-\alpha/2}}{2}\right]^2 \qquad \cdots\cdots(12)$$

式中：$u_{1-\alpha/2}$ 是标准正态分布的 $1-\alpha/2$ 分位数。

c_2 是满足下式的最小整数：

$$c_2\geqslant\left[\sqrt{n\lambda_0+\frac{4-u_{1-\alpha/2}^2}{12}}+\frac{u_{1-\alpha/2}}{2}\right]^2 \qquad \cdots\cdots(13)$$

5.3 单侧检验 $H_0:\lambda\leqslant\lambda_0$ $H_1:\lambda>\lambda_0$

5.3.1 实施步骤

a) 由 λ_0，样本量 n 及给定的显著性水平 α，确定拒绝域的临界值 c_2（c_2 的确定见 5.3.2）。

b) 计算 $T=\sum_{i=1}^{n}x_i$ 的值。

c) 当 $T\geqslant c_2$ 时，拒绝 H_0；当 $T<c_2$ 时，不拒绝 H_0。

5.3.2 拒绝域临界值 c_2 的确定

c_2 可由 λ_0,n 及 α 确定。它是满足下式的最小整数:

$$P\{T \geqslant c_2 \mid n\lambda_0\} = \sum_{k=c_2}^{\infty} \frac{(n\lambda_0)^k}{k\,!} e^{-n\lambda_0} \leqslant \alpha$$

相应的第一类错误的概率 α' 和第二类错误的概率 β' 见附录 A。c_2 可由 χ^2 分布表法或正态近似法确定,优先采用 χ^2 分布表法,当有困难时,可采用正态近似法。

5.3.2.1 χ^2 分布表法

c_2 是满足下式的最小整数:

$$2n\lambda_0 \leqslant \chi^2_{\alpha}(2c_2) \qquad \cdots\cdots\cdots\cdots (14)$$

式中 $\chi^2_{\alpha}(2c_2)$ 是自由度为 $2c_2$ 的 χ^2 分布的 α 分位数。

5.3.2.2 正态近似法

c_2 是满足下式的最小整数:

$$c_2 \geqslant \left[\sqrt{n\lambda_0 + \frac{4-u^2_{1-\alpha}}{12}} + \frac{u_{1-\alpha}}{2}\right]^2 \qquad \cdots\cdots\cdots\cdots (15)$$

式中:$u_{1-\alpha}$ 是标准正态分布的 $1-\alpha$ 分位数。

同双侧检验一样,当利用 χ^2 分布表法有困难时,可使用正态近似法。

5.4 单侧检验 H_0:$\lambda \geqslant \lambda_0$ H_1:$\lambda < \lambda_0$

5.4.1 实施步骤

a) 由 λ_0、样本量 n 及给定的显著性水平 α,确定拒绝域的临界值 c_1(c_1 的确定见 5.4.2)。

b) 计算 $T = \sum_{i=1}^{n} x_i$ 的值。

c) 当 $T \leqslant c_1$ 时,拒绝 H_0;当 $T > c_1$ 时,不拒绝 H_0。

5.4.2 拒绝域临界值 c_1 的确定

c_1 由 λ_0、n 和 α 确定,它是满足下式的最大整数。

$$P\{T \leqslant c_1 \mid n\lambda_0\} = \sum_{k=0}^{c_1} \frac{(n\lambda_0)^k}{k\,!} e^{-n\lambda_0} \leqslant \alpha$$

相应的第一类错误的概率和第二类错误的概率见附录 A。c_1 可由 χ^2 分布表法或正态近似法确定,优先采用 χ^2 分布表法。

5.4.2.1 χ^2 分布表法

c_1 是满足下式的最大整数:

$$2n\lambda_0 \geqslant \chi^2_{1-\alpha}(2c_1 + 2) \qquad \cdots\cdots\cdots\cdots (16)$$

式中 $\chi^2_{1-\alpha}(2c_1+2)$ 是自由度为 $2c_1+2$ 的 χ^2 分布的 $1-\alpha$ 分位数。

5.4.2.2 正态近似法

c_1 是满足下式的最大整数:

$$c_1 + 1 \leqslant \left[\sqrt{n\lambda_0 + \frac{4-u^2_{1-\alpha}}{12}} - \frac{u_{1-\alpha}}{2}\right]^2 \qquad \cdots\cdots\cdots\cdots (17)$$

式中 $u_{1-\alpha}$ 是标准正态分布的 $1-\alpha$ 分位数。

优先采用 χ^2 分布表法,当利用 χ^2 分布表法有困难时,可使用正态近似法。

附　录　A
（规范性附录）
显著性检验的两类错误

A.1　第一类错误

第一类错误的概率是指当原假设为真时，错误拒绝原假设的概率。它可表示为拒绝域的临界值的函数。记作 α'。

原假设 $H_0:\lambda=\lambda_0$，备择假设 $H_1:\lambda\neq\lambda_0$ 情形：

$$\begin{aligned}\alpha' &= P(T\leqslant c_1 \mid n\lambda_0)+P(T\geqslant c_2 \mid n\lambda_0)\\ &= P(T\leqslant c_1 \mid n\lambda_0)+1-P(T\leqslant c_2-1 \mid n\lambda_0)\end{aligned}$$

原假设 $H_0:\lambda\leqslant\lambda_0$，备择假设 $H_1:\lambda>\lambda_0$ 情形：

$$\alpha' = P(T\geqslant c_2 \mid n\lambda_0) = 1-P(T\leqslant c_2-1 \mid n\lambda_0)$$

原假设 $H_0:\lambda\geqslant\lambda_0$，备择假设 $H_1:\lambda<\lambda_0$ 情形：

$$\alpha' = P(T\leqslant c_1 \mid n\lambda_0)$$

由于泊松分布的离散性，拒绝域的临界值是整数，α'常常不是正好等于给定的显著性水平 α，而是比 α 小。

A.2　第二类错误

第二类错误的概率是原假设非真时，未拒绝原假设的概率，它可以表示为拒绝域的临界值和所考虑的备择假设中特定的 λ 值的函数，记作 β'。

原假设 $H_0:\lambda=\lambda_0$ 情形，备择假设 $H_1:\lambda\neq\lambda_0$ 情形：$\beta'=P(c_1<T<c_2|n\lambda)$

$$=P(T\leqslant c_2-1|n\lambda)-P(T\leqslant c_1|n\lambda)$$

原假设 $H_0:\lambda\leqslant\lambda_0$ 情形，备择假设 $H_1:\lambda>\lambda_0$ 情形：$\beta'=P(T\leqslant c_2-1)|n\lambda)$

原假设 $H_0:\lambda\geqslant\lambda_0$ 情形，备择假设 $H_1:\lambda<\lambda_0$ 情形：$\beta'=P(T>c_1|n\lambda)=1-P(T\leqslant c_1|n\lambda)$

β'表示真实的第二类错误的概率，这个记号的使用仿效第一类错误的情况，检验的功效（即原假设不成立时，拒绝它的概率）为 $1-\beta'$。

A.3　第一类错误的概率 α'和第二类错误的概率 β'的计算

在给定 H_0、α 和 n 的情况下，在确定了拒绝域的临界值 c_1 和 c_2 后，可进一步计算：

a)　实际的第一类错误的概率 α'是多大？

b)　对应于特别指定的备择假设 $H_1:\lambda=\lambda_1$，第二类错误的概率 β'是多大？

c)　指定第二类错误的概率为 β，其对应的 λ 值是多大？

表 A.1 给出了计算 α'和 β'的几种方法。

问题 c)的解决见表 A.2。

表 A.1 中的正态近似方法，只适用于 c_1 和 c_2 较大时（大于 15）。对 $H_0:\lambda=\lambda_0$ 情形，如果 c_2 较大，而 c_1 较小，在计算 α'或 β'时，可对 $P(T\leqslant c_2-1|n\lambda_0)$或 $P(T\leqslant c_2-1|n\lambda_1)$采用正态近似，而对 $P(T\leqslant c_1-1|n\lambda_0)$或 $P(T\leqslant c_1-1|n\lambda_1)$利用 χ^2 分布分位数表计算。

A.4　例

对 5.2.2.1 中的例：

原假设 $H_0:\lambda=\lambda_0=0.6$，$n=15$，显著性水平 $\alpha=0.10$，确定出的拒绝域的临界值 $c_1=3$，$c_2=15$。15 个样本观测值的总和 $T=\sum_{i=1}^{15}x_i=18$，进一步计算：