



运筹学

讲者： 顾乃杰 教授、 黄章进 副教授



排队论

Chap. 13 *Queueing Theory*



13. 排队论

- 13.1 基本概念
- 13.2 到达间隔的分布和服务时间的分布
- 13.3 单服务台负指数分布排队系统的分析
- 13.4 多服务台负指数分布排队系统的分析
- 13.5 一般服务时间 **M/G/1** 模型
- 13.6 经济分析—系统的最优化

13.5 一般服务时间M/G/1模型

- 服务时间是任意分布的情形

– 任何情形下列关系都成立：

$$E[\text{系统中顾客数}] = E[\text{队列中顾客数}] + E[\text{服务机构中顾客数}]$$

$$E[\text{在系统中逗留时间}] = E[\text{排队等候时间}] + E[\text{服务时间}]$$

其中 $E[\cdot]$ 表示求期望值，用符号表示即：

$$L_s = L_q + L_{se}, \quad W_s = W_q + E[T] \quad (13.37)$$

T 表示服务时间（随机变量），当 T 服从负指数分布时， $E[T] = 1/\mu$ 。

– 结合Little公式中的 $L_s = \lambda W_s, \quad L_q = \lambda W_q$

- 上述七个指标中只要知道3个就可求出其余的指标；
- 在有限源或者队长有限制的情况下， λ 要换成有效到达率 λ_{eff} 。

13.5 一般服务时间M/G/1模型

- Pollaczek-Khintchine(P-K)公式

- 对于M/G/1模型，服务时间 T 的分布是一般的（但要求期望值 $E[T]$ 和方差 $Var[T]$ 都存在），其他条件和标准的M/M/1型相同。为了达到稳态，要求 $\rho = \lambda E[T] < 1$ 。满足上述条件下有：

$$L_s = \rho + \frac{\rho^2 + \lambda^2 Var[T]}{2(1 - \rho)} \quad (13.38)$$

Pollaczek-Khintchine(P-K) 公式

- 不管 T 是什么分布，只要知道 λ ， $E[T]$ 和 $Var[T]$ ，就可以求出 L_s ，并根据各项指标之间的关系式求出 L_q 、 W_q 和 W_s 。
- 需要注意的是因为有方差项的存在，所以若在研究各期望值时不考虑概率性质，会得出错误结果，仅当 $Var[T]=0$ 时，随机性的波动才不会影响 L_s 。要想改进各指标，除考虑期望值外，还可以从改变方差来考虑。

13.5 一般服务时间M/G/1模型

- 例9：有一售票口，已知顾客按平均为2分30秒的时间间隔的负指数分布到达，顾客在售票口前服务时间平均为2分钟。在下面两种情况下求顾客购票的平均逗留时间和等待时间：(1)服务时间也服从负指数分布，求顾客为购票所需平均逗留时间和等待时间；(2)经过调查，顾客在售票窗口至少要占用1分钟，服从以下概率密度分布：

$$f(y) = \begin{cases} e^{-y+1}, & y \geq 1 \\ 0, & y < 1 \end{cases}$$

- 分析：

(1) 是典型的M/M/1模型， $\lambda = 1/2.5 = 0.4$ ， $\mu = 1/2 = 0.5$ ，

$$\rho = \lambda / \mu = 0.8, \quad W_s = \frac{1}{\mu - \lambda} = 10 \text{分}, \quad W_q = \frac{\rho}{\mu - \lambda} = 8 \text{分}$$

(2) 令y为服务时间，那么 $y = 1 + x$ ，代入概率密度函数，得：

$$f(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad \text{可见} x \text{服从均值为1的负指数分布，则：}$$

$$E[y] = 2, \quad \text{Var}[y] = \text{Var}[1 + x] = \text{Var}[x] = 1, \quad \rho = \lambda E[y] = 0.8,$$

$$\text{代入} P-K \text{公式得：} L_s = 0.8 + \frac{0.8^2 + 0.4^2 \times 1}{2 \times (1 - 0.8)} = 2.8, \quad L_q = L_s - \rho = 2,$$

$$W_s = L_s / \lambda = 7 \text{分},$$

$$W_q = L_q / \lambda = 5 \text{分}$$

13.5 一般服务时间M/G/1模型

- 定长服务时间M/D/1模型

- 服务时间是确定的常数，此时：

$$L_s = \rho + \frac{\rho^2 + \lambda^2 \text{Var}[T]}{2(1-\rho)} \quad (13.38)$$

$$T = 1/\mu, \quad \text{Var}[T] = 0, \quad L_s = \rho + \frac{\rho^2}{2(1-\rho)} \quad (13.39)$$

- 例10：某实验室有一台自动检验机器性能的仪器，检验机器的顾客按泊松分布到达，每小时平均4个顾客，检验每台机器需要6分钟。

- 分析： $\lambda = 4$, $E[T] = 1/10$ (小时), $\rho = 4/10$, $\text{Var}[T] = 0$

- 在检验室内机器台数期望值： $L_s = 0.4 + \frac{(0.4)^2}{2(1-0.4)} = 0.533$ (台)

- 等候检验的机器台数期望值： $L_q = 0.533 - 0.4 = 0.133$ (台)

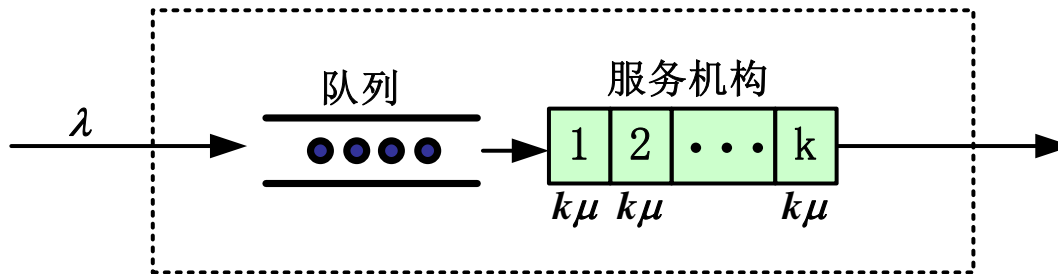
- 每台机器在室内逗留时间期望值： $W_s = \frac{0.533}{4} = 0.133$ 小时

- 每台机器平均等待检验时间期望值： $W_q = \frac{0.133}{4} = 0.033$ 小时

13.5 一般服务时间M/G/1模型

• 爱尔朗服务时间 $M/E_k/1$ 模型

- 如果顾客必须经过 k 个服务站，在每个服务站的服务时间 T_i 相互独立并服从相同的负指数分布（参数为 $k\mu$ ），则： $T = \sum_{i=1}^k T_i$ 服从 k 阶爱尔朗分布。



$$E[T_i] = \frac{1}{k\mu}$$

$$Var[T_i] = \frac{1}{k^2\mu^2}$$

$$E[T] = \frac{1}{\mu}$$

$$Var[T] = \frac{1}{k\mu^2}$$

- 代入P-K公式，该模型的各项指标如下：

$$\left. \begin{aligned} L_s &= \rho + \frac{\rho^2 + \frac{\lambda^2}{k\mu^2}}{2(1-\rho)} = \rho + \frac{(k+1)\rho^2}{2k(1-\rho)}, & L_q &= \frac{(k+1)\rho^2}{2k(1-\rho)} \\ W_s &= L_s / \lambda, & W_q &= L_q / \lambda \end{aligned} \right\} \quad (13.40)$$

13.5 一般服务时间M/G/1模型

- 例11：某单人裁缝店做西服，每套需经过4个不同的工序，4个工序完成后才能开始做另一套。每一个工序的时间服从负指数分布，期望值为2小时。顾客到来服从泊松分布，平均订货率为5.5套/周（设一周6天为工作日，每天8小时）。顾客为等到做好一套西服的期望时间有多长？
- 分析： $\lambda = 5.5$ 套/周， $1/(4\mu) = 2$ 小时， $\mu = \frac{1}{8}$ 套/小时 = 6套/周， $\rho = 5.5/6$

$$E[T_i] = 2, \quad Var[T_i] = \left(\frac{1}{4 \times 6}\right)^2$$

$$E[T] = 8, \quad Var[T] = \frac{1}{4 \times 6^2}$$

$$L_s = \frac{5.5}{6} + \frac{\left(\frac{5.5}{6}\right)^2 + (5.5)^2 \times \frac{1}{4 \times 6^2}}{2\left(1 - \frac{5.5}{6}\right)} = 7.2188$$

顾客为等到做好一套西服的期望时间： $W_s = L_s / \lambda = 1.3$ 周

13.6 经济分析——系统的最优化

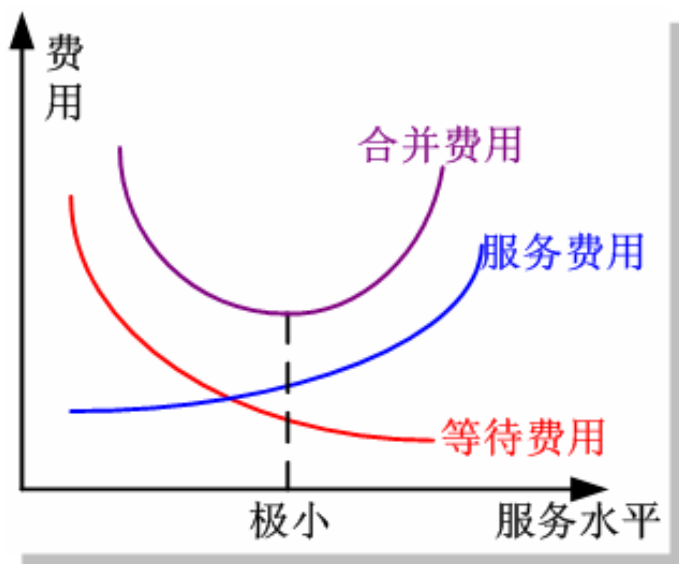
- 排队系统的最优化问题

- 系统设计的最优化

- 称为静态问题，目的在于使设备达到最大效益，在一定的质量指标下要求机构最为经济。

- 系统控制的最优化

- 称为动态问题，是指给定一个系统，如何运营可使某个目标函数得到最优。



□ 在一般情形下，提供服务水平（数量、质量）自然会降低顾客的等待费用（损失），但却常常增加了服务机构的成本。

□ 最优化的目标之一是使二者费用之和为最小；另一个常用的目标函数是使纯收入或使利润（服务收入与服务成本之差）最大。



13.6 经济分析——系统的最优化

— 费用

- 各种费用在稳态情形下都是按单位时间来考虑的，一般情形，服务费用（成本）可以确切计算或估计，但是顾客的等待费用（损失）有许多不同情况。

— 服务水平

- 平均服务率 μ （代表服务机构的服务能力和经验等）；
- 服务设备，如服务台个数 c ；
- 由队列所占空间大小所决定的队列最大限制数 N ；
- 服务水平也可以通过服务强度 ρ 来表示。

— 常用的求解方法：

- 对于离散变量常用边际分析法；
- 对于连续变量常用经典的微分法；
- 对于复杂的问题，也可以使用非线性规划或动态规划方法。

13.6 经济分析——系统的最优化

- M/M/1模型中最优服务率 μ

- 标准的M/M/1模型

取目标函数 z 为单位时间服务成本与 顾客在系统逗留费用之 和的期望值:

$$z = c_s \mu + c_w L_s \quad (13.41)$$

其中 c_s 为当 $\mu = 1$ 时服务机构单位时间的 费用,

c_w 为每个顾客在系统停留 单位时间的费用。

又 $L_s = \frac{\lambda}{\mu - \lambda}$, 因此 $z = c_s \mu + c_w \cdot \frac{\lambda}{\mu - \lambda}$, 求极小值,

令: $\frac{dz}{d\mu} = c_s - c_w \cdot \frac{\lambda}{(\mu - \lambda)^2} = 0$, 解得:

$$\mu^* = \lambda + \sqrt{\frac{c_w}{c_s} \lambda} \quad (\text{为了保证 } \rho < 1, \mu > \lambda, \text{ 根号前取 } + \text{ 号})$$

13.6 经济分析——系统的最优化

— 系统中顾客最大限制数为N的情形

- 系统中若有N个顾客，则后来的顾客被拒绝。 P_N 即为被拒绝的概率（呼损率）， $1-P_N$ 即为能接受服务的概率， $\lambda(1-P_N)$ 即为单位时间实际进入服务机构顾客的平均数（有效到达率），在稳定状态下，等于单位时间内实际服务完成的平均顾客数。
- 解：

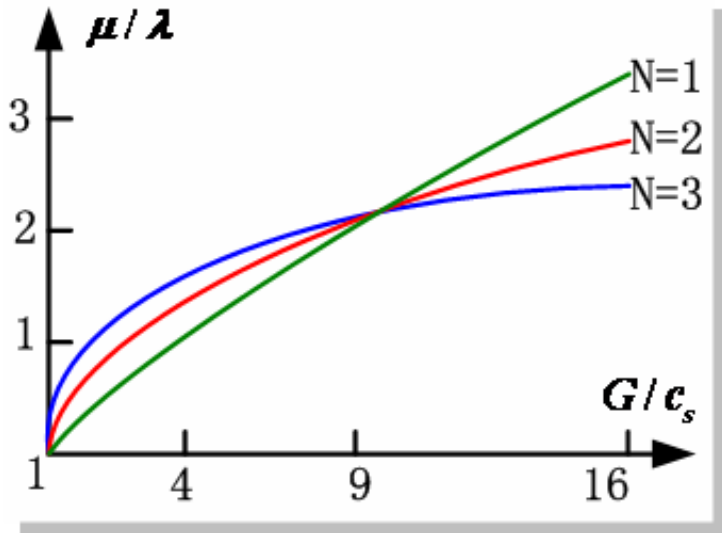
设每服务 1 人能收入 G 元，于是单位时间收入的期望值是 $\lambda(1-P_n)G$ 元，纯利润是：

$$z = \lambda(1-P_n)G - c_s \mu$$

$$= \lambda G \cdot \frac{1-\rho^N}{1-\rho^{N+1}} - c_s \mu = \lambda \mu G \cdot \frac{\mu^N - \lambda^N}{\mu^{N+1} - \lambda^{N+1}} - c_s \mu$$

$$\text{令 } \frac{dz}{d\mu} = 0, \text{ 得 } \rho^{N+1} \cdot \frac{N - (N+1)\rho + \rho^{N+1}}{(1-\rho^{N+1})^2} = \frac{c_s}{G}$$

最优的 μ^* 应满足上式，其中 c_s 、 G 、 λ 、 N 都是给定的，通常通过数值计算来求出 μ^* ，或将上式左方（对一定的 N ）作为 ρ 的函数做出图形，对于给定的 G/c_s ，根据图形可求出 μ^*/λ 。



13.6 经济分析——系统的最优化

— 顾客源为有限的情形（以机器故障问题为例）

- 共有机器 m 台以及1个修理工人，各台连续运转时间和修理时间均服从负指数分布。当服务率 $\mu=1$ 的时的修理费用 c_s ，单位时间每台机器运转可得收入 G 元。平均运转台数为 $m-L_s$ ，所以单位时间纯利润为：

$$z = (m - L_s)G - c_s \mu = \frac{mG}{\rho} \cdot \frac{E_{m-1}\left(\frac{m}{\rho}\right)}{E_m\left(\frac{m}{\rho}\right)} - c_s \mu$$

其中 $E_m(x) = \sum_{k=0}^m \frac{x^k}{k!} e^{-x}$ 称为泊松部分和；

令 $\rho = \frac{m\lambda}{\mu}$ 代入上式，得：

$$z = \frac{mG}{\rho} \cdot \frac{E_{m-1}\left(\frac{m}{\rho}\right)}{E_m\left(\frac{m}{\rho}\right)} - c_s \frac{m\lambda}{\rho}$$

如何得到？

13.6 经济分析——系统的最优化

—由 (13-28) 和 (13-27) 式知道:

$$L_s = m - \frac{\mu}{\lambda}(1 - P_0)$$

$$P_0 = \frac{1}{\sum_{i=0}^m \frac{m!}{(m-i)!} \left(\frac{\lambda}{\mu}\right)^i}$$

令 $\rho = \frac{m\lambda}{\mu}$, $E_m(x) = \sum_{k=0}^m \frac{x^k}{k!} e^{-x}$ 得:

$$\begin{aligned} z &= (m - L_s)G - c_s \mu = \frac{\mu}{\lambda}(1 - P_0)G - c_s \mu = \frac{mG}{\rho} \left(1 - \frac{1}{\sum_{i=0}^m \frac{m!}{(m-i)!} \left(\frac{\lambda}{\mu}\right)^i}\right) - c_s \mu \\ &= \frac{mG}{\rho} \cdot \frac{\sum_{i=1}^m \frac{m!}{(m-i)!} \left(\frac{\lambda}{\mu}\right)^i}{\sum_{i=0}^m \frac{m!}{(m-i)!} \left(\frac{\lambda}{\mu}\right)^i} - c_s \mu = \frac{mG}{\rho} \cdot \frac{\sum_{i=1}^m \frac{1}{(m-i)!} \left(\frac{\lambda}{\mu}\right)^i}{\sum_{i=0}^m \frac{1}{(m-i)!} \left(\frac{\lambda}{\mu}\right)^i} - c_s \mu = \frac{mG}{\rho} \cdot \frac{\sum_{k=0}^{m-1} \frac{1}{k!} \left(\frac{\lambda}{\mu}\right)^{m-k}}{\sum_{k=0}^m \frac{1}{k!} \left(\frac{\lambda}{\mu}\right)^{m-k}} - c_s \mu \\ &= \frac{mG}{\rho} \cdot \frac{\sum_{k=0}^{m-1} \frac{1}{k!} \left(\frac{\lambda}{\mu}\right)^{-k}}{\sum_{k=0}^m \frac{1}{k!} \left(\frac{\lambda}{\mu}\right)^{-k}} - c_s \mu = \boxed{\frac{mG}{\rho} \cdot \frac{E_{m-1}\left(\frac{m}{\rho}\right)}{E_m\left(\frac{m}{\rho}\right)} - c_s \frac{m\lambda}{\rho}} \end{aligned}$$

13.6 经济分析——系统的最优化

— 顾客源为有限的情形（以机器故障问题为例）

- 共有机器 m 台以及1个修理工人，各台连续运转时间和修理时间均服从负指数分布。当服务率 $\mu=1$ 的时的修理费用 c_s ，单位时间每台机器运转可得收入 G 元。平均运转台数为 $m-L_s$ ，所以单位时间纯利润为：

$$z = (m - L_s)G - c_s \mu = \frac{mG}{\rho} \cdot \frac{E_{m-1}\left(\frac{m}{\rho}\right)}{E_m\left(\frac{m}{\rho}\right)} - c_s \mu$$

其中 $E_m(x) = \sum_{k=0}^m \frac{x^k}{k!} e^{-x}$ 称为泊松部分和；

令 $\rho = \frac{m\lambda}{\mu}$ 代入上式，得：

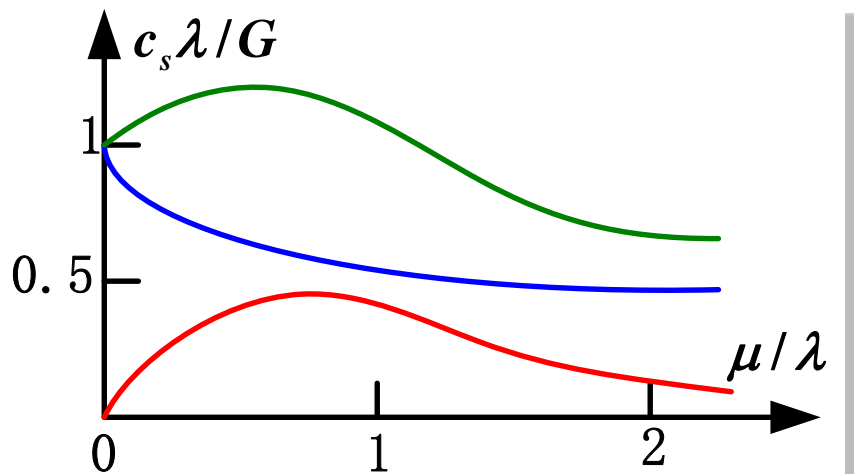
$$z = \frac{mG}{\rho} \cdot \frac{E_{m-1}\left(\frac{m}{\rho}\right)}{E_m\left(\frac{m}{\rho}\right)} - c_s \frac{m\lambda}{\rho}$$

13.6 经济分析——系统的最优化

为了求得最优的 μ^* ，令 $\frac{dz}{dx} = 0$ ，整体对 $\frac{m}{\rho}$ 求导

利用 $\frac{d}{dx} E_m(x) = E_{m-1}(x) - E_m(x)$ ，得：

$$\frac{E_{m-1}\left(\frac{m}{\rho}\right)E_m\left(\frac{m}{\rho}\right) + \frac{m}{\rho} \left[E_m\left(\frac{m}{\rho}\right)E_{m-2}\left(\frac{m}{\rho}\right) - E_{m-1}^2\left(\frac{m}{\rho}\right) \right]}{E_m^2\left(\frac{m}{\rho}\right)} = \frac{c_s \lambda}{G}$$



给定 m 、 G 、 c_s 、 λ ，求解 μ^* 很困难，通常利用泊松分布表通过数值计算获得，或将上式左方（对一定的 m ）作为 ρ 的函数做出图形，对于给定的 $\frac{c_s \lambda}{G}$ 根据图形可求出 μ^* / λ 。



13.6 经济分析——系统的最优化

- M/M/c 模型中最优的服务台数 c
 - 标准的M/M/c模型在稳态情形下单位时间全部费用（服务成本与等待费用之和）的期望值：

$$z = c'_s \cdot c + c_w \cdot L \quad (13.43)$$

- 其中

c : 服务台数;

c'_s : 每服务台单位时间的成本;

c_w : 为每个顾客在系统停留单位时间的费用;

L : 系统中顾客平均数 L_s 或 队列中等待的顾客平均数 L_q

（它们随 z 值的不同而不同）

因为 c'_s 和 c_w 都是给定的，唯一可能变动的是服务台数 c ，所以 z 是 c 的函数 $z(c)$ 。

13.6 经济分析——系统的最优化

- 因为 c 只取整数值， $z(c)$ 不是连续变量的函数，采用边际分析法 (Marginal Analysis)，根据 $z(c^*)$ 最小的特点：

$$\begin{cases} z(c^*) \leq z(c^* - 1) \\ z(c^*) \leq z(c^* + 1) \end{cases} \Rightarrow \begin{cases} c'_s \cdot c^* + c_w L(c^*) \leq c'_s (c^* - 1) + c_w L(c^* - 1) \\ c'_s \cdot c^* + c_w L(c^*) \leq c'_s (c^* + 1) + c_w L(c^* + 1) \end{cases}$$

$$\Rightarrow L(c^*) - L(c^* + 1) \leq c'_s / c_w \leq L(c^* - 1) - L(c^*)$$

- 依次求 $c = 1, 2, 3, \dots$ 时 L 的值，并作两相邻的 L 值之差，因 c'_s / c_w 是已知数，根据这个数落在哪个不等式区间里就可以确定出 c^* 。

13.6 经济分析——系统的最优化

- 例12：某检验中心为各工厂服务，要求做检验的工厂的到来服从泊松流，平均到达率 λ 为每天48次，每次来检验由于停工等原因损失为6元。服务时间服从负指数分布，平均服务率 μ 为每天25次，每设置一个检验员服务成本为每4元，其他条件符合标准 M/M/c 模型。问应设置几个检验员才能使总费用的期望值最小？
- 分析： $c'_s = 4$ 元/每检验员， $c_w = 6$ 元/次， $\lambda = 48$ ， $\mu = 25$ ， $\lambda / \mu = 1.92$
设检验员数为 c ，令 c 依次为 1、2、3、4、5，根据已有的计算表求出 L_s ：

| c | 1 | 2 | 3 | 4 | 5 |
|---|------|---------|--------|--------|--------|
| $\lambda / c\mu$ | 1.92 | 0.96 | 0.64 | 0.48 | 0.38 |
| 查多服务台的 $W_q \cdot \mu$ 计算表 | - | 10.2550 | 0.3961 | 0.0772 | 0.0170 |
| $L_s = \frac{\lambda}{\mu} (W_q \cdot \mu + 1)$ | - | 21.610 | 2.680 | 2.068 | 1.952 |

13.6 经济分析——系统的最优化

– 将 L_s 代入得表：

| 检验员数 c | 来检验顾客数 $L_s(c)$ | $L(c)-L(c+1) \sim$ $L(c)-L(c-1)$ | 总费用(每天) $z(c)$ |
|-------------|--------------------|-------------------------------------|-------------------|
| 1 | ∞ | | ∞ |
| 2 | 21.610 | 18.930 ~ ∞ | 154.94 |
| 3 | 2.680 | 0.612 ~ 18.930 | 27.87 |
| 4 | 2.068 | 0.116 ~ 0.612 | 28.38 |
| 5 | 1.952 | | 31.71 |

$\frac{c'_s}{c_w} = 0.66$ ，落在区间 $(0.612 \sim 18.930)$ 内

所以 $c^* = 3$

即设 3 个检验员使总费用最小，此时 $z(c^*) = 27.87$ (元)

本章完

The end