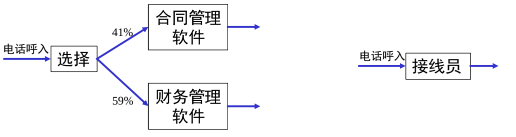
五. 仿真输出估计比较与精度控制

1. 系统方案设计比较
2. 方差衰减技术
3. 区间估计方法比较
4. 仿真实验比较

例. SMP公司呼叫中心有7名接线员，负责解答顾客的问题，自动化系统会自动选择对应的产品线。目前财务软件4名接线员，合同软件3名接线员。SMP公司想知道是否可以通过技能交叉培训合并产品线，从而减少接线员总人数。如果交叉培训，则接线员处理每名顾客的时间会增加10%。



评价指标：平均服务时间，顾客呼入到问题解决的平均时间

1. 系统设计方案比较

由于方案输出指标可能互不独立、样本量与精度各不相同，直接比较均值不符合统计学要求

两方案比较
- (1) 配对检验法
- (2) Welch法

多方案比较
- 标准比较法

两方案比较

(1) 配对检验法

存在两个方案A_1、A_2,仿真随机观测数列$X_{11},X_{12},\ldots,X_{1n_1}$和$X_{21},X_{22},\ldots,X_{2n_2}$，其中$n_1$和$n_2$不必相等，$\mu_1$和$\mu_2$分别是两个方案的数学期望。

当$n_1=n_2$时，令$n=n_1=n_2$;当$n_1\neq n_2$时，令$n=min\{n_1,n_2\}$。定义$Z_j=X_{1j}-X_{2j}$，则$Z_1,Z_2,\ldots,Z_n$是独立同分布的随机变量

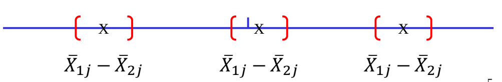
(1) 配对检验法

$$\bar{Z}(n)=\frac{\sum_{j=1}^{n}Z_j}{n}$$

$$S^2(n)=\frac{\sum_{j=1}^{n}[Z_j-\bar{Z}(n)]^2}{n-1}$$

$100(1-\alpha)\%$置信水平下的置信区间为

$$\bar{Z}(n)\pm t_{1-\frac{\alpha}{2}}(n-1)S(n)/\sqrt{n}$$



(1) 配对检验法

- 若置信区间包含零值，可认为使用当前的验证法，无法获得X_{1j}与X_{2j}存在显著差异的结论；
- 若置信区间不包含零值，且置信区间的上下界均小于零，可认为X_{1j}与X_{2j}存在显著差异，X_{1j}整体上小于X_{2j}；
- 若置信区间不包含零值，且置信区间的上下界均大于零，可认为X_{1j}与X_{2j}存在显著差异，X_{1j}整体上大于X_{2j}。

(1) Welch法

存在两个方案A_1、A_2,仿真随机观测数列$X_{11}, X_{12}, \dots, X_{1n_1}$和$X_{21}, X_{22}, \dots, X_{2n_2}$，其中$n_1 \neq n_2$，假设$X_{1j}$与$X_{2j}$相互独立且服从正态分布。

$$\bar{X}_1 = \frac{\sum_{j=1}^{n_1} X_{1j}}{n_1}, \bar{X}_2 = \frac{\sum_{j=1}^{n_2} X_{2j}}{n_2}$$

$$S_1^2(n) = \frac{\sum_{j=1}^{n_1}[X_{1j}-\bar{X}_1]^2}{n_1-1}, S_2^2(n) = \frac{\sum_{j=1}^{n_2}[X_{2j}-\bar{X}_2]^2}{n_2-1}$$

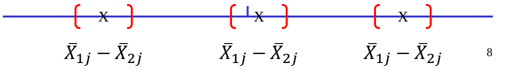
(1) Welch法

$100(1-\alpha)\%$置信水平下的置信区间为

$$\bar{X}_1(n_1) - \bar{X}_2(n_2) \pm t_{1-\frac{\alpha}{2}}(\hat{f})\sqrt{\frac{S_1^2(n_1)}{n_1} + \frac{S_2^2(n_2)}{n_2}}$$

其中，自由度的估计值为

$$\hat{f} = \frac{[S_1^2(n_1)/n_1 + S_2^2(n_2)/n_2]^2}{[S_1^2(n_1)/n_1]^2/(n_1-1) + [S_2^2(n_2)/n_2]^2/(n_2-1)}$$



多方案比较

对于多个方案的比较，仍需使用置信区间法。为了满足总体置信水平，需要使用Bonferroni不等式限定每个置信区间的置信水平。

当总体置信水平设定为$1-\alpha$时，如果要求c个置信区间有相同的置信水平，则置信水平应限定在$1-\alpha/c$。

标准比较法

① 标准方案编号为1，其他方案顺序编号2, 3,…,k

② 采用配对检验采用配对检验法或Welch法，可以构成k-1个置信区间，分别基于$\mu_2-\mu_1, \mu_3-\mu_1, \dots, \mu_k-\mu_1$，具有$1-\alpha/(k-1)$的置信水平。

③ 分别研究每个区间是否包含零值，若不包含则说明方案与标准方案有差异；若包含则认为该方案与基准方案没有显著差异。

④ 经过筛选，从k-1个比较方案中选出数量不等的优异方案

⑤ 从中设定一个基准方案，重复以上步骤，最终找到最优方案。

例 考虑SMP公司的三种方案，$R_0 = 10$，总体置信水平95%

方案C：现有系统

方案P7：交叉培训系统，7名接线员

方案P6：交叉培训系统，6名接线员

需构建$\theta_C - \theta_{P7}$，$\theta_{P7} - \theta_{P6}$，$\theta_C - \theta_{P6}$三个置信区间，总误差概率0.05，令$\alpha_i = 0.05/3 = 0.0167$，$t_{1-\frac{\alpha_i}{2}}(9) = 2.97$，则在总体95%置信水平下，

$$\theta_C - \theta_{P7}: -0.86 \pm 1.84$$

$$\theta_C - \theta_{P6}: -5.91 \pm 4.22$$

$$\theta_{P7} - \theta_{P6}: -5.05 \pm 2.61$$

说明方案C和方案P7优于方案P6

	平均响应时间			差异	
序号	方案C	方案P7	方案P6	C-P7	P7-P6
1	6.24	6.19	8.35	0.05	-2.14
2	9.06	10.64	18.03	-1.59	-7.39
3	8.02	9.53	16.17	-1.51	-6.64
4	5.93	6.15	7.40	-0.22	-1.25
5	8.31	7.83	12.70	0.48	-4.87
6	5.91	6.09	8.26	-0.17	-2.17
7	8.74	7.62	12.32	1.12	-4.70
8	7.78	7.03	11.40	0.75	-4.37
9	7.15	12.79	23.04	-5.64	-10.24
10	5.72	7.57	14.30	-1.85	-6.73
样本均值	7.29	8.14	13.20	-0.86	-5.05
样本方差	1.60	4.86	23.99	3.87	7.72
标准误差				0.62	0.88

2. 方差衰减技术

仿真精度
- 绝对精度：$\theta = t_{1-\alpha/2}(R-1) \cdot S / \sqrt{R}$
- 相对精度：$\phi = \dfrac{t_{1-\alpha/2}(R-1) \cdot S / \sqrt{R}}{\bar{Y}}$

仿真的精度是用仿真输出样本的方差加以评估的，为提高精度，必须采用适当的方差衰减技术。

(1) 公共随机数法

主要用于比较两个不同系统模型的差异。通过采用相同的随机输入，尽可能消除由于随机因素造成的仿真输出差异，而使这种差异主要表现为系统模型本身的差异。

原理

设两个系统的输出变量分别为X_1与X_2，预估计$\mu = E(X_1)-E(X_2)$。

对两模型分别进行n次独立重复仿真运行，得到观测值分别为X_{1i}与X_{2i} ($i=1,2,\cdots,n$)。

令$X_i = X_{1i} - X_{2i}$，$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$，则有

$$E(\bar{X}_n) = \mu$$

$$Var(\bar{X}_n) = \frac{1}{n}\left[Var(X_1) + Var(X_2) - 2Cov(X_1, X_2)\right]$$

$$\begin{aligned} Var(\bar{X}_n) &= Var\left\{\frac{1}{n}\sum_{i=1}^{n}(X_{1i} - X_{2i})\right\} \\ &= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_{1i} - X_{2i}) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_1 - X_2) \\ &= \frac{1}{n}\left[Var(X_1) + Var(X_2) - 2Cov(X_1, X_2)\right] \end{aligned}$$

对两个系统的相应输入使用公共随机数，希望使X_{1i}与X_{2i}正相关，从而减小方差$Var(\bar{X}_n)$。

实施方法

对不同系统模型中的相同特征的输入随机变量，使用同一随机数发生器和相同的种子值。

(2) 对偶变量法

主要用于一个系统仿真运行的方差衰减，通过利用互补随机数对系统进行成对仿真运行（对某一随机变量，第一次仿真用u_k生成，而第二次仿真则用$1-u_k$生成），尽可能消除随机因素对仿真结果的影响。

原理

利用互补随机数分别进行n对独立重复仿真运行，得到观测值分别为$(X_{1i},\ X_{2i})$ $(i=1,2,\cdots,n)$，现估计 $\mu = E(X)$。

令$X_i = (X_{1i} + X_{2i})/2$

则 $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ 是$\mu = E(X)$的无偏估计。

由于$X_i(i=1,2,\cdots,n)$相互独立，易得

$$Var(\bar{X}_n) = Var\left\{\frac{1}{2n}\sum_{i=1}^{n} X_{1i} + X_{2i}\right\}$$
$$= \frac{1}{4n^2}\sum_{i=1}^{n} Var(X_{1i} + X_{2i})$$
$$= \frac{1}{4n^2}\sum_{i=1}^{n} Var(X_{1i}) + Var(X_{2i}) + 2Cov(X_{1i}, X_{2i})$$

由互补随机数的取法希望X_{1i}与 X_{2i}负相关，从而减小方差$Var(\bar{X}_n)$。

(3) 控制变量法

主要用于同一系统仿真运行的方差衰减，是通过利用随机变量的相关性实现的。

原理

设X为系统的输出随机变量，欲估计$\mu = E(X)$，又设Y为与X相关的另一随机变量，且$E(Y)=\nu$为已知。令$X_c = X - a(Y - \nu)$，易知有

$$E(X_c) = E(X) = \mu$$

适当选取a，可得

$$Var(X_c) = Var(X) - \frac{[Cov(Y, X)]^2}{Var(Y)}$$

只要X, Y相关，即可实现方差衰减，相关性越强，衰减越明显。称Y为X的控制变量。

实施方法

通常可取系统中某输入随机变量(分布为已知)作为控制变量，这样，控制变量就不需额外生成了。

例如，对于排队系统，若输出变量X为平均等待时间，则可取控制变量Y为到达时间间隔（与X负相关），或服务时间（与X正相关）。

五. 仿真输出估计比较与精度控制

3. 区间估计方法比较

仿真输出中的一个重要问题是建立有关系统位置参数的置信区间估计。点估计可以给出未知参数的一个好的推测，而区间估计说明了这个推测的误差多大合理。

当感兴趣的参数是总体均值时，最常用的方法是用独立重复或批平均值法收集仿真输出数据，然后使用参数统计技术，如z或t统计量建立置信区间。当感兴趣的位置参数是总体中位数时，可以利用样本中位数的大样本性质进行区间估计。

4

设θ是一个感兴趣的系统中心位置参数，而 $X_1, X_2, \ldots, X_n$ 是 n 次独立仿真运行的输出值，它们构成总体X的一个样本。我们的目的是利用仿真输出数据获得系统位置参数θ的置信区间估计。

(1) 方法1(经典方法)步骤：

1) 计算点估计(样本均值): $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

2) 计算方差估计(样本方差): $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$

3) 对给定显著水平α，计算均值θ的近似100(1－α)%置信区间为

$$(\bar{X}_n - t_{\alpha/2}(n-1)\cdot S/\sqrt{n},\ \bar{X}_n + t_{\alpha/2}(n-1)\cdot S/\sqrt{n})$$

(2) 方法2

当总体分布的概率密度关于未知点$\theta(-\infty<\theta<+\infty)$是一个对称函数，则对称中心$\theta$既是总体未知分布的均值(如果均值存在)，又是它的一个中位数。因而可以引出θ的两个估计：样本均值 $\bar{X}_n$ 和样本中位数 $\tilde{X}_n$。若总体均值存在，则样本中位数也是θ的无偏估计；若密度函数$f(x)$在中位数θ处连续，且$f(\theta)>0$，则当$n\to\infty$时，统计量

$$\sqrt{n}(\tilde{X}_n - \theta)$$

渐近服从正态分布

$$N(0,\ 1/[4f^2(\theta)])$$

步骤：

1) 计算点估计(样本中位数):

$$\tilde{X}_n = \begin{cases} X_{((n+1)/2)}, & n\text{是奇数}, \\ \dfrac{X_{(n/2)} + X_{(n/2+1)}}{2}, & n\text{是偶数}. \end{cases}$$

其中 $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$ 是仿真输出的顺序统计量；

2) 计算方差估计: $Var(\tilde{X}_n) \approx \frac{1}{4nf^2(\theta)}$

3) 对给定显著水平α，计算均值θ的近似100(1－α)%置信区间为

$$\left(\tilde{X}_n - \frac{z_{\alpha/2}}{2\sqrt{n}\,f(\tilde{X}_n)},\ \tilde{X}_n + \frac{z_{\alpha/2}}{2\sqrt{n}\,f(\tilde{X}_n)}\right)$$

(3) 方法3 (Bootstrap方法)

通常仿真输出数据的分布是未知的。当总体分布未知，需要确定某个估计量的抽样分布时，Bootstrap方法是一个有效的选择。Bootstrap方法是通过从原始样本中进行有放回抽样来建立一个估计量的经验分布(称之为这个估计的Bootstrap分布)，它就是所求抽样分布的一个近似。这个分布完全以原始样本为基础，不依赖任何理论分布，于是可以象抽样分布那样用Bootstrap分布建立系统参数的置信区间。

假定 $t(X_1, X_2, \ldots, X_n)$ 是参数θ的估计量，用Bootstrap方法建立θ的置信区间的步骤为

1) 从原始样本中进行有放回抽样，抽取m个独立的容量为n的Bootstrap样本：

$$X_1^i, X_2^i, \ldots, X_n^i,\ i = 1, 2, \ldots, m$$

2) 对每个i (i=1, 2, …, m), 计算

$$t(X_1^i, X_2^i, \ldots, X_n^i)$$

的值；

3) 用2)中获得的m个值为 t 建立一个Bootstrap分布；

4) 利用 t 的Bootstrap分布及其百分位数建立θ的置信区间。

Bootstrap方法应用范围：

- 服从任何分布样本
- 任何样本量
- 任何感兴趣的估计量。