统计质量管理作业

sword

2019-09-22

目录

1 第一周作业

1. 说明下面三个抽样方案有类似的 OC 曲线。

(i) 一次计数抽样方案：$(20\,,1)$；

(ii) 二次计数抽样方案：$n_1=n_2=13$，判定数组为 $(0\,,2\,,1\,,2)$；

(iii) 五次计数抽样方案：$n_1=n_2=\cdots=n_5=5$，判定数组为 $(\#\,,2\,,0\,,2\,,0\,,2\,,0\,,2\,,0\,,2\,,1\,,2)$。

解答

(i) 在一次计数抽样方案中，接收概率为

$$AR_1(p)=P\{X_1\le 1\}=\sum_{i=0}^{1}C_{20}^{i}p^{i}(1-p)^{20-i}$$

其中 p 为不合格率，下同。

(ii) 在二次计数抽样方案中，接收概率为

$$\begin{aligned}AR_2(p)&=P\{X_1=0\}+P\{X_1=1\,,X_1+X_2\le 1\}\\&=(1-p)^{13}+C_{13}^{1}p(1-p)^{12}\times(1-p)^{13}\\&=(1-p)^{13}+13p(1-p)^{25}\end{aligned}$$

(iii) 在五次计数抽样方案中，接收概率为

$$\begin{aligned}AR_3(p)&=P\{\sum_{i=1}^{2}X_i=0\}+P\{\sum_{i=1}^{2}X_i=1\,,\sum_{i=1}^{3}X_i=1\,,\sum_{i=1}^{4}X_i=1\,,\sum_{i=1}^{5}X_i=1\}\\&=(1-p)^{10}+C_{10}^{1}p(1-p)^{9}\times(1-p)^{5}\times(1-p)^{5}\times(1-p)^{5}\\&=(1-p)^{10}+10p(1-p)^{24}\end{aligned}$$

于是，令 p 在 $[0\,,1]$ 之间取不同的值，分别计算接收概率，这里借助 Python 软件算出具体数值并绘图如下。

表 1: 不同抽样方式下的接收概率

p	0	0.01	0.05	0.1	0.2	1.0
$AR_1(p)$	1.000	0.983	0.736	0.392	0.069	0.000
$AR_2(p)$	1.000	0.979	0.694	0.348	0.068	0.000
$AR_3(p)$	1.000	0.983	0.745	0.428	0.117	0.000

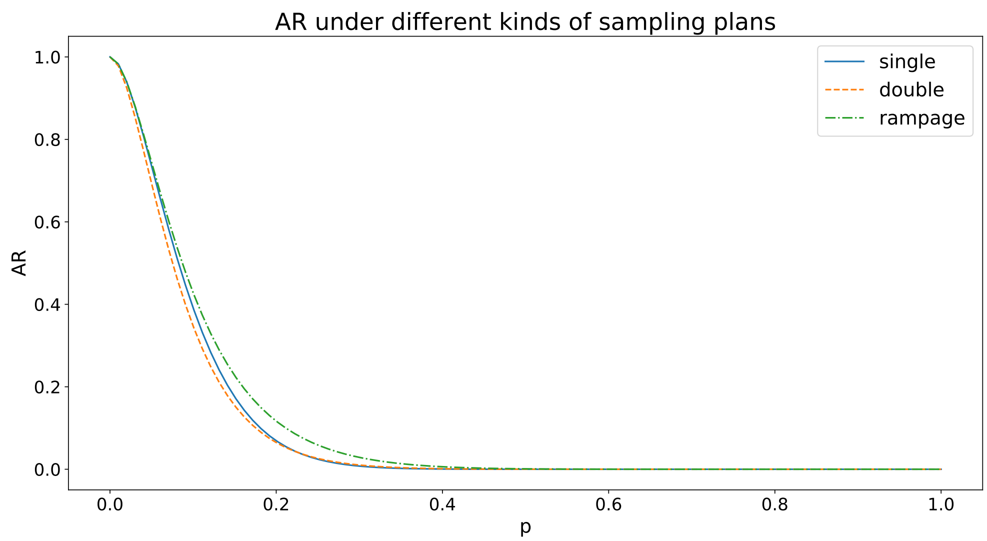

图 1: 不同抽样方式下的接收概率

从图中可以看出，三条 OC 曲线均呈现下凸的递减态势。

附：具体代码如下（环境：Python 3.7.2 & vscode 1.37.1）

```
from scipy import stats
import numpy as np
import matplotlib.pyplot as plt

p = np.linspace(0, 1, 100)
ar1 = stats.binom.cdf(1, 20, p)
ar2 = (1-p)**13 + 13*p*(1-p)**25
ar3 = (1-p)**10 + 10*p*(1-p)**24
plt.plot(p, ar1, linestyle="-", label="single")
plt.plot(p, ar2, linestyle="--", label="double")
plt.plot(p, ar3, linestyle="-.", label="rampage")
plt.legend(loc="best", fontsize=18)
plt.title("AR under different kinds of sampling plans", fontsize=22)
plt.xlabel("p", fontsize=18)
plt.ylabel("AR", fontsize=18)
plt.xticks(fontsize=16)
plt.yticks(fontsize=16)
plt.show()
```

2 第二周作业

1. 某电子仪器厂与协作的电容器厂商定，当电容器厂提供的产品批的不合格品率不超过 3% 时以高于 95% 的概率接收，而当不合格品率超过 12% 时，将以低于 10% 的概率接收。现有一批总计 2000 件产品，试制定计数标准型一次抽样检验方案，画出 OC 曲线，并与上例进行比较，对二者结果有差异给出解释。附上程序。

解答

记一次抽检方案为 $(n\,,c)$，它表示在这批产品中随机抽取 n 个，若不合格品不多于 c 个则接收，否则拒绝。记 p 为不合格率，随机变量 $X \sim \mathrm{h}(n\,,2000\,,2000p)$，则接收概率为

$$\begin{aligned} L(p) &= P\{X \le c\} \\ &= \frac{\binom{2000p}{c}\binom{2000-2000p}{n-c}}{\binom{2000}{n}} \end{aligned}$$

据题设，有

$$\begin{aligned} L(p) > 95\% &\qquad p \le 3\% \\ L(p) < 10\% &\qquad p > 12\% \end{aligned}$$

这等价于

$$\begin{aligned} L(0.03) > 0.95 \\ L(0.12) < 0.10 \end{aligned}$$

于是，借助 Python 软件搜索满足上式的 $(n\,,c)$，部分结果如下。

表 2: 2000 件产品的一次抽检方案部分结果

n	65	66	67	75	$\cdots$	121	121	121	122	$\cdots$	2000
c	4	4	4	5	$\cdots$	7	8	9	7	$\cdots$	239

为节约成本，取 $(n\,,c) = (65\,,4)$。借助 Python 软件绘出 OC 曲线如下。

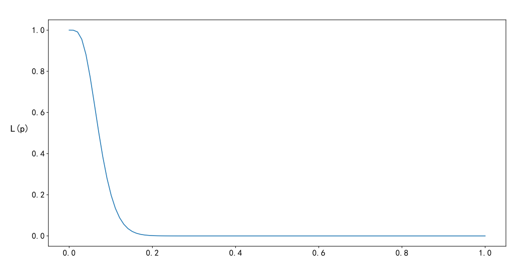

图 2: $(65\,,4)$ 下的 OC 曲线

上例的结果为 $(66\,,4)$ 。这说明，在有限总体情形下，检验所需样品的个数可以减少，而效果相同。事实上，总体有限的情形对应不放回抽样，总体无限的情形对应有放回抽样。抽取相同样品个数的情形下，不放回的包含的信息更多。进一步的试验显示，总体数为 100 个时，只需抽取 47 个样品。

2. 对一批钢材的强度进行抽样检验，按质量要求，规定强度大于 42 $\mathrm{kg/mm^2}$ 为合格品，已知强度服从正态分布，其标准差为 4 $\mathrm{kg/mm^2}$ 。现在生产方与使用方商定，若不合格品率不超过 3% 时以不低于 95% 的概率接收，而当不合格品率超过 12% 时，将以不高于 10% 的概率接收。试制定计量标准型一次抽样检验方案。

解答

记强度所服从正态分布的均值为 μ ，从这批钢材中随机抽出 n 件样品，平均强度为 $\bar{x}$ 。于是接收概率为

$$\begin{aligned} L(\mu) &= P\{\bar{x} \geq k_L\} \\ &= P\left\{\frac{\sqrt{n}(\bar{x}-\mu)}{4} \geq \frac{\sqrt{n}(k_L-\mu)}{4}\right\} \\ &= 1-\Phi\left(\frac{\sqrt{n}(k_L-\mu)}{4}\right) \end{aligned}$$

其中，k_L 为待定常数。

设随机变量 $X \sim \mathrm{N}(\mu\,,16)$ ，则不合格率 p 为

$$\begin{aligned} p &= P\{X < 42\} \\ &= P\left\{\frac{X-\mu}{4} < \frac{42-\mu}{4}\right\} \\ &= \Phi\left(\frac{42-\mu}{4}\right) \end{aligned}$$

不合格品率不超过 3% 时，有

$$\begin{aligned} p \leq 3\% &\Rightarrow \Phi\left(\frac{42-\mu}{4}\right) \leq 3\% \\ &\Rightarrow \frac{42-\mu}{4} \leq u_{0.03} \\ &\Rightarrow \mu \geq 42-4u_{0.03} \end{aligned}$$

同样地，不合格品率超过 12% 时，有 $\mu < 42-4u_{0.12}$ 。从而，据题设，得

$$\begin{aligned} 1-\Phi\left(\frac{\sqrt{n}(k_L-\mu)}{4}\right) &\geq 95\% \qquad \mu \geq 42-4u_{0.03} \\ 1-\Phi\left(\frac{\sqrt{n}(k_L-\mu)}{4}\right) &\leq 10\% \qquad \mu < 42-4u_{0.12} \end{aligned}$$

这等价于

$$1-\Phi\left(\frac{\sqrt{n}(k_L-42+4u_{0.03})}{4}\right)\geq 0.95$$
$$1-\Phi\left(\frac{\sqrt{n}(k_L-42+4u_{0.12})}{4}\right)\leq 0.10$$

亦即

$$\frac{\sqrt{n}(k_L-42+4u_{0.03})}{4}\leq u_{0.05}$$
$$\frac{\sqrt{n}(k_L-42+4u_{0.12})}{4}\geq u_{0.90}$$

此处我们取边界条件，即

$$\frac{\sqrt{n}(k_L-42+4u_{0.03})}{4}=u_{0.05}$$
$$\frac{\sqrt{n}(k_L-42+4u_{0.12})}{4}=u_{0.90}$$

由此解出

$$\begin{aligned}k_L&=\frac{42u_{0.90}-4u_{0.03}u_{0.90}-42u_{0.05}+4u_{0.12}u_{0.05}}{u_{0.90}-u_{0.05}}\\&=47.94\end{aligned}$$

继而得到

$$\begin{aligned}n&\leq\left(\frac{4u_{0.05}}{k_L-42+4u_{0.03}}\right)^2\\&=17.27\end{aligned}$$

取 $n=17$ 。综上，我们得到一次抽检方案为 $(17, 47.94)$ ，即从这批钢材中随机抽取 17 件样品，若平均强度不小于 47.94 $\mathrm{kg/mm^2}$ 则接收，否则拒绝。

附：具体实现代码如下（环境：Python 3.7.2 & vscode 1.37.1）

(i) 计算抽检方案

```
from scipy import stats
import numpy as np

p0 = 0.03
p1 = 0.12
a = 0.95
b = 0.10
N = 2000
n_all = []
```

```
c_all = []
for n in range(1, N+1):
    for c in range(1, n+1):
        l0 = stats.hypergeom.cdf(c, N, N*p0, n)
        l1 = stats.hypergeom.cdf(c, N, N*p1, n)
        if l0 > a and l1 < b:
            n_all.append(n)
            c_all.append(c)
```

(ii) 绘制 OC 曲线

```
from scipy import stats
import numpy as np
import matplotlib.pyplot as plt

p = np.linspace(0, 1, 100)
l = stats.hypergeom.cdf(4, 2000, 2000*p, 65)
plt.plot(p, l)
plt.xlabel("p", fontsize=18)
plt.ylabel("L(p)", fontsize=18,
           rotation="horizontal", horizontalalignment="right")
plt.xticks(fontsize=16)
plt.yticks(fontsize=16)
plt.show()
```

3 第三周作业

1. 对一批钢材的强度进行抽样检验，按质量要求，规定强度大于 42 $\mathrm{kg/mm^2}$ 为合格品，已知强度服从正态分布，其标准差为 4 $\mathrm{kg/mm^2}$ 。现在生产方与使用方商定，若不合格品率不超过 3% 时以不低于 95% 的概率接收，而当不合格品率超过 12% 时，将以不高于 10% 的概率接收。试制定计量型序贯抽样检验方案。

解答

记强度所服从正态分布的均值为 μ ，$\alpha = 0.05$ ，$\beta = 0.10$ 。我们知道 AQL=3% 对应 $\mu \geq \mu_0 = 42 - 4u_{0.03}$ ，LQL=12% 对应 $\mu < \mu_1 = 42 - 4u_{0.12}$ 。于是，此检验等价于

$$H_0 : \mu \geq \mu_0 \qquad \text{v.s.} \qquad H_1 : \mu < \mu_1$$

对抽得的样本 $(x_1, x_2, \cdots, x_n)$ ，定义逐次检验比 SPR_n 为

$$SPR_n = \prod_{i=1}^{n} \frac{\varphi\left(\frac{x_i-\mu_1}{\sigma}\right)}{\varphi\left(\frac{x_i-\mu_0}{\sigma}\right)}$$

其中，$\varphi(\cdot)$ 为标准正态分布的密度函数。
当 H_0 成立时，x_i 集中于 μ_0 附近而远离 μ_1 ，因此上式中分母大、分子小，SPR_n 偏小；反之，H_1 成立时，SPR_n 偏大。由此，我们需要确定两个常数 A, B ，当 $SPR_n \geq B$ 时拒绝原假设，$SPR_n \leq A$ 时不拒绝原假设，$A < SPR_n < B$ 时继续抽样。对 A 和 B ，Wald 给出的建议是

$$A = \frac{\beta}{1-\alpha} \qquad B = \frac{1-\beta}{\alpha}$$

综上，我们得到了如下的抽检方案。

表 3: 钢材强度检验的序贯型方案

$SPR_n \leq \frac{\beta}{1-\alpha}$	接收这批钢材
$SPR_n \geq \frac{1-\beta}{\alpha}$	拒收这批钢材
$\frac{\beta}{1-\alpha} < SPR_n < \frac{1-\beta}{\alpha}$	继续抽样

进一步，对 SPR_n 取对数，有

$$\ln SPR_n = \frac{\mu_1 - \mu_0}{2\sigma^2} \sum_{i=1}^{n} (2x_i - \mu_0 - \mu_1)$$

从而对上述方案的表达式进行简化，代入具体数值，得到如下表格。

表 4: 钢材强度检验的序贯型方案（简化形式）

$\sum_{i=1}^{n} x_i \leq 48.11n - 16.38$	接收这批钢材
$\sum_{i=1}^{n} x_i \geq 48.11n + 12.76$	拒收这批钢材
$48.11n - 16.38 < \sum_{i=1}^{n} x_i < 48.11n + 12.76$	继续抽样

2. 设批量 $N = 1000$ ，试求一次计数抽样方案 $(20\,,1)$ 在 $p = 0\,,0.01\,,0.05\,,0.10$ 处的接收概率。

解答

易见接收概率为

$$L(p) = P\{h(1000\,,1000p\,,20) \leq 1\}$$

计算得结果如下。

表 5: 一次方案 $(20\,,1)$ 的接收概率

p	0	0.01	0.05	0.10
$L(p)$	1.00	0.98	0.74	0.39

3. 设产品的批量 $N = 5000$ ，采用一次抽样方案 $(20\,,1)$ 。
(1) 求 $p = 1\%\,,5\%\,,10\%$ 时的平均抽检量；
(2) 若长期采用这一方案，求平均检出质量。

解答

(1) 平均抽检量 ATI 为

$$\mathrm{ATI} = nL(p) + N(1 - L(p)) = N - (N - n)L(p)$$

其中 $L(p)$ 为接收概率。从而计算出在不同不合格率 p 下的平均抽检量如下。

表 6: 不同不合格率下的平均抽检量

p	1%	5%	10%
ATI	103	1335	3052

(2) 平均检出质量（AOQ）为经过检查后产品的平均不合格率，在检查中发现的不合格品均须用合格品代替。

假设检验的总批数为 K ，定义随机变量

$$Y_i = \begin{cases} 1 & \text{接收第 } i \text{ 批产品} \\ 0 & \text{拒收第 } i \text{ 批产品} \end{cases} \qquad i = 1\,, 2\,, \cdots\,, K$$

易见 $Y_i \overset{\text{i.i.d.}}{\sim} b(1\,, L(p))$ 。

在这 K 批产品中，被接收的批数为 $\sum_{i=1}^{K} Y_i$ ，且被接收的产品批中抽检出的不合格品均由合格品代替。因此检出质量为

$$\frac{\sum_{i=1}^{K} Y_i (N-n)p}{KN}$$

于是，长期来看，平均检出质量为

$$\begin{aligned} &\lim_{K\to\infty} \mathrm{E}\left(\frac{\sum_{i=1}^{K} Y_i (N-n)p}{KN}\right) \\ =&\frac{N-n}{N} pL(p) \\ =&\frac{N-n}{N} pP\{h(N\,, Np\,, n) \leq 1\} \end{aligned}$$

借助 Python 软件绘出平均检出质量随不合格率变动的图像如下。

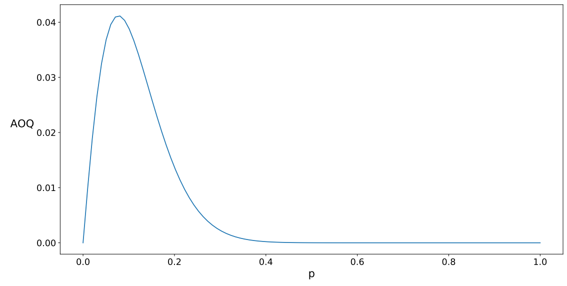

图 3: 方案 $(20\,, 1)$ 下的平均检出质量

4. 某厂生产的一种产品中要求氧化铁的含量要低，批均值不超过 0.004（%）以高于 0.95 的概率接收，批均值不低于 0.005（%）时只能以低于 0.10 的概率接收，已知该含量服从正态分布，且 $\sigma = 0.0006$ （%），试给出计量标准性一次抽样方案。

解答

记批均值为 μ ，$\mu_1 = 0.004\%\,, \mu_2 = 0.005\%\,, \alpha = 0.05\,, \beta = 0.10$ 。设满足题意的抽样方案为 $(n\,, k_L)$ ，即从一批产品中抽出 n 件，若氧化铁含量的均值 $\bar{x}$ 不高于 k_L 则接收，否则拒收。于是，接收概率为

$$L(\mu) = P\{\bar{x} \le k_L\} = \Phi\left(\frac{k_L - \mu}{\sigma/\sqrt{n}}\right)$$

据题设，有

$$\begin{aligned} L(\mu) &> 1 - \alpha \qquad \mu \le \mu_1 \\ L(\mu) &< \beta \qquad \mu \le \mu_2 \end{aligned}$$

由于 $L(\mu)$ 关于 μ 单调递减，故上式等价于

$$\begin{aligned} L(\mu_1) &> 1 - \alpha \\ L(\mu_2) &< \beta \end{aligned}$$

此处，取边界条件，就有

$$\begin{aligned} \frac{k_L - \mu_1}{\sigma/\sqrt{n}} &> u_{1-\alpha} \\ \frac{k_L - \mu_2}{\sigma/\sqrt{n}} &< u_\beta \end{aligned}$$

由此解出 $k_L = 0.00687(\%)\,, n = 34$ 。
因此，一次抽样方案可以为 $(34\,, 0.00687(\%))$ ，即从一批产品中抽出 34 件，若氧化铁含量的均值不高于 0.00687（%）则接收，否则拒收。

5. 现有一批产品需验收，规定 $p_0 = 0.018\,, p_1 = 0.18\,, \alpha = 0.05\,, \beta = 0.10$ ，试给出符合此规定的计数序贯抽检方案及其图形。

解答

设在某次抽样过程中抽取到的产品序列为 $(x_1\,, x_2\,, \cdots\,, x_n)$ 且前 $n-1$ 次抽样并未得出结果，其中诸 x_i 或者为 0（合格），或者为 1（不合格）。记

$$
\begin{aligned}
w &= \ln \frac{p_1(1-p_0)}{p_0(1-p_1)} = 2.483 \\
s &= -\ln \frac{1-p_1}{1-p_0} / w = 0.073 \\
d_1 &= \ln \frac{\beta}{1-\alpha} / w = -0.907 \\
d_2 &= \ln \frac{1-\beta}{\alpha} / w = 1.164
\end{aligned}
$$

则符合题意的计数序贯抽检方案如下。

表 7: 某产品的计数序贯抽检方案

$\sum_{i=1}^{n} x_i \leq 0.073n - 0.907$	接收这批产品
$\sum_{i=1}^{n} x_i \geq 0.073n + 1.164$	拒收这批产品
$0.073n - 0.907 < \sum_{i=1}^{n} x_i < 0.073n + 1.164$	继续抽样

借助 Excel 软件绘出图形如下。

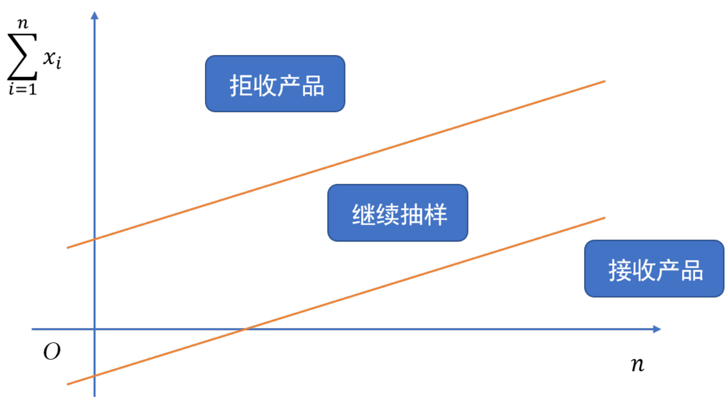

图 4: 某产品的计数序贯抽检方案量

附：绘制平均检出质量图的代码如下（环境：Python 3.7.2 & vscode 1.37.1）

```
import numpy as np
from scipy import stats
import matplotlib.pyplot as plt


N = 5000
n = 20
```

```
p = np.linspace(0, 1, 100)
AOQ = (N-n)*p*stats.hypergeom.cdf(1, N, N*p, n)/N
plt.plot(p, AOQ)
plt.xlabel("p", fontsize=18)
plt.ylabel("AOQ", fontsize=18, rotation="horizontal",
           horizontalalignment="right")
plt.xticks(fontsize=15)
plt.yticks(fontsize=15)
plt.show()
```