Stat/IE 531 Solutions to Homework #5

1. (Problem 4.5 of the notes)
With $I = 10, J = 5, K = 500$, the expected mean squares are

   - $EMSA = KJ\sigma_\alpha^2 + K\sigma_\beta^2 + \sigma^2$,
   - $EMSB(A) = K\sigma_\beta^2 + \sigma^2$, and
   - $EMSC(B(A)) = \sigma^2$.

   (a) The expected value of the grand sample variance can be written as

$$\begin{aligned}
Es^2 &= \frac{1}{IJK-1}\sum(y_{ijk} - \bar{y}_{...})^2 \\
&= \frac{1}{IJK-1} ESS_{TOTAL} \\
&= \frac{1}{IJK-1}\{ESS_A + ESS_{B(A)} + ESS_{C(B(A))}\} \\
&= \frac{1}{IJK-1}\{(I-1)(KJ\sigma_\alpha^2 + K\sigma_\beta^2 + \sigma^2) + I(J-1)(K\sigma_\beta^2 + \sigma^2) + IJ(K-1)\sigma^2\} \\
&= .9\sigma_\alpha^2 + .98\sigma_\beta^2 + \sigma^2.
\end{aligned}$$

   (b) From SAS output, $\sigma_\alpha^2 = 14.4232$, $\sigma_\beta^2 = .142$, $\sigma^2 = .484$.
Hence, $Es^2 = .9(14.4232) + .98(.142) + .484 = 13.604$, and
$E[s^2|\sigma_\alpha^2 = 0] = .98(.142) + .484 = .62316$.
(SAS OUTPUT)

```
                          Type 1 Analysis of Variance

                                                 Sum of
Source                     DF                   Squares          Mean Square
week                        4                580.000000           145.000000
day(week)                  20                 15.360000             0.768000
Error                      25                 12.100000             0.484000
Corrected Total            49                607.460000                    .


                          Type 1 Analysis of Variance
Source                    Expected Mean Square
week                      Var(Error) + 2 Var(day(week)) + 10 Var(week)
day(week)                 Var(Error) + 2 Var(day(week))
Error                     Var(Error)
Corrected Total           .
                                 Type 1 Estimates
                       Variance Component          Estimate
                       Var(week)                   14.42320
                       Var(day(week))               0.14200
                       Var(Error)                   0.48400
```

2. (Problem 4.6 of the notes)

(a) Scenario 1: Each crate contains 250 widgets with diameter 5 and 250 widgets with diameter 7

- Population

Source	df	SS	MS	EMS
Crates	99	0	0	$\sigma^2 + 500\sigma^2_{crate}$
Widgets	49900	50000	1.002	σ^2
Total	49999	50000	1.00002	

$$\Rightarrow \qquad \sigma^2_{widget} = 1.00002, \qquad \sigma^2_{crate} = -.002 \approx 0$$

- Sample (4 crates, 5 widgets per crate)

Source	df	EMS
Crates	3	$\sigma^2 + 5\sigma^2_{crate}$
Widgets	16	σ^2_{widget}
Total	19	

Scenario 2: All crates are diameter 5 or 7, but no mixtures

Source	df	SS	MS	EMS
Crates	99	50000	505.05	$\sigma^2 + 500\sigma^2_{crate}$
Widgets	49900	0	0	σ^2
Total	49999	50000		

$$\Rightarrow \qquad \sigma^2_{widget} = 0, \qquad \sigma^2_{crate} = 1.0101\cdots$$

The expected value of the sample variance of the 20 widgets:

$$\begin{aligned}
E\hat{\sigma}^2 &= E\left(\frac{1}{19}\sum_{i=1}^{4}\sum_{j=1}^{5}(y_{ij}-\bar{y}_{..})^2\right)\\
&= \frac{1}{19}E\left(5\sum_{i=1}^{4}(\bar{y}_{i.}-\bar{y}_{..})^2 + \sum_{i=1}^{4}\sum_{j=1}^{5}(y_{ij}-\bar{y}_{i.})^2\right)\\
&= \frac{1}{19}E\left(3\left[\frac{5\sum(\bar{y}_{i.}-\bar{y}_{..})^2}{3}\right] + 16\left[\frac{\sum\sum(y_{ij}-\bar{y}_{i.})^2}{16}\right]\right)\\
&= \frac{1}{19}[3EMS_{crates} + 16EMS_{widgets}]\\
&= \frac{1}{19}[3(5\sigma^2_{crate} + \sigma^2_{widget}) + 16\sigma^2_{widget}]\\
&= \frac{15}{19}\sigma^2_{crate} + \sigma^2_{widget}
\end{aligned}$$

For Scenario 1, $E\hat{\sigma}^2 = \frac{15}{19}(0) + 1.00002 = 1.00002$.

For Scenario 2, $E\hat{\sigma}^2 = \frac{15}{19}(1.0101\cdots) + 0 = .797448$.

(b) Use the formula on page 51 of the notes:

$$\frac{M(N-1)}{NM-1}s_A^2 + \left(\frac{N(M-1)}{NM-1} - \frac{M(N-1)}{NM-1}\left(\frac{1}{m} - \frac{1}{M}\right)\right)s_B^2$$

with $N = 100, M = 500, m = 5$,

$$s_A^2 = (1/3)\sum(\bar{y}_{i.} - \bar{y}_{..})^2$$

$$s_B^2 = (1/4)(s_1^2 + s_2^2 + s_3^2 + s_4^2), \qquad \text{where } s_i^2 = \textstyle\sum_{j=1}^5 (y_{ij} - \bar{y}_{i.})^2/4$$

$= .99s_A^2 + .802s_B^2$

3. (Problem 4.7 of the notes)

(a) $ER_i = d_2(4)\sigma \quad \Rightarrow \quad \widehat{\sigma} = \bar{R}/d_2(4) = .00435/2.059 = .0021127$

$\Rightarrow \quad \widehat{\sigma}^2 = \underline{4.46 \times 10^{-6}}$

$E\Delta_i = d_2(10)\sqrt{\sigma_\alpha^2 + \sigma^2/4}$

Now $\overline{\text{range}} = .0155 - (-.012250) = .02775$.

Solving $\quad .02775 = 3.078\sqrt{\widehat{\sigma}_\alpha^2 + (4.46 \times 10^{-6})/4} \quad$ for $\widehat{\sigma}_\alpha^2$,

we get $\widehat{\sigma}_\alpha^2 = \underline{8.012 \times 10^{-5}}$.

Based on ANOVA table from SAS output, we have $\widehat{\sigma}_\alpha^2 = .00008232$, $\widehat{\sigma}^2 = 4.66 \times 10^{-6}$.

```
                          Type 1 Analysis of Variance
                                             Sum of
       Source                     DF        Squares     Mean Square
       bundle                      9       0.003005        0.000334
       Error                      30       0.000140     0.000004656
       Corrected Total            39       0.003145               .


                          Type 1 Analysis of Variance
      Source                Expected Mean Square
      bundle                Var(Error) + 4 Var(bundle)
      Error                 Var(Error)
      Corrected Total       .
                                Type 1 Estimates
                      Variance Component          Estimate
                      Var(bundle)               0.00008232
                      Var(Error)                4.65625E-6
```

(b) Use the fact that $\widehat{\sigma}_\alpha^2 = \frac{1}{4}[MS_{bundle} - MS_{error}]$ to compute $\sqrt{\text{Var}(\widehat{\sigma}_\alpha^2)}$, i.e.,

$$\begin{aligned}\sqrt{\text{Var}(\widehat{\sigma}_\alpha^2)} &= \sqrt{2\left[\left(\frac{1}{4}\right)^2\frac{.00033392^2}{9} + \left(\frac{1}{4}\right)^2\frac{.00000466^2}{30}\right]} \\ &= \sqrt{2[4.71 \times 10^{-11} + 4.52408 \times 10^{-14}]} \\ &= 3.935 \times 10^{-5}\end{aligned}$$

(c) $.00008232 \pm 1.645(3.935 \times 10^{-5}) = (1.76 \times 10^{-5}, 1.47 \times 10^{-4})$.

4. (Problem 5.1 of the notes)

(a) Percent defective

$$P_a(p) = \sum_{x=0}^{c} \frac{\begin{pmatrix} Np \\ x \end{pmatrix} \begin{pmatrix} N(1-p) \\ n-x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}$$

Mean defects

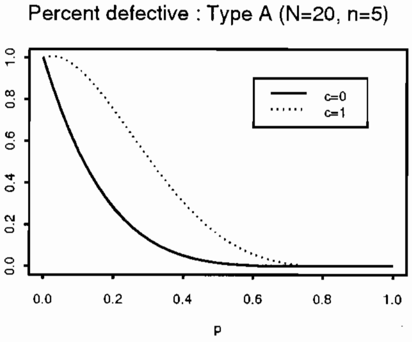
$$P_a(\lambda) = \sum_{x=0}^{c} \begin{pmatrix} N\lambda \\ x \end{pmatrix} \left(\frac{n}{N}\right)^x \left(1 - \frac{n}{N}\right)^{N\lambda - x}$$

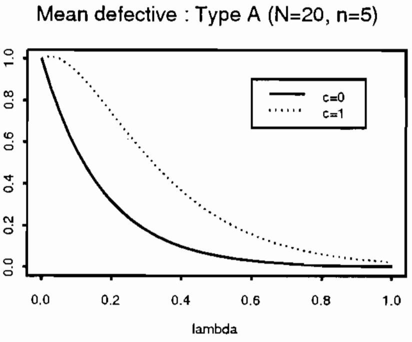
(b) Percent defective

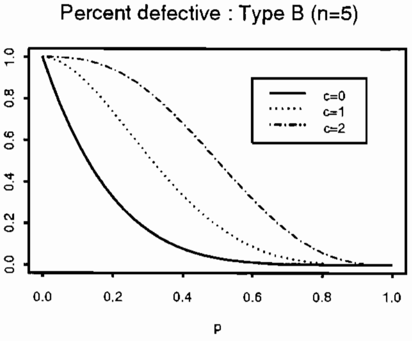
$$P_a(p) = \sum_{x=0}^{c} \begin{pmatrix} n \\ x \end{pmatrix} p^x (1-p)^{n-x}$$

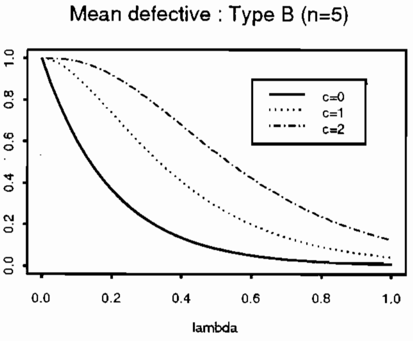
Mean defects

$$P_a(\lambda) = \sum_{x=0}^{c} \frac{\exp(-n\lambda)(n\lambda)^x}{x!}$$


Percent defective : Type A (N=20, n=5)


Mean defective : Type A (N=20, n=5)


Percent defective : Type B (n=5)


Mean defective : Type B (n=5)

5. (Problem 5.2 of the notes)

(a) We want approximate OC curves for (1) $n = 100, C = 1$, (2) $n = 200, C = 2$, (3) $n = 300, C = 3$:

(1) $P_a(p) = \begin{pmatrix} 100 \\ 0 \end{pmatrix} p^0(1-p)^{100} + \begin{pmatrix} 100 \\ 1 \end{pmatrix} p^1(1-p)^{99} = (1-p)^{100} + 100p(1-p)^{99}$

$$(2)\ P_a(p) = \begin{pmatrix} 200 \\ 0 \end{pmatrix} p^0(1-p)^{200} + \begin{pmatrix} 200 \\ 1 \end{pmatrix} p^1(1-p)^{199} + \begin{pmatrix} 200 \\ 2 \end{pmatrix} p^2(1-p)^{198}$$
$$= (1-p)^{200} + 200p(1-p)^{199} + 19900p^2(1-p)^198$$

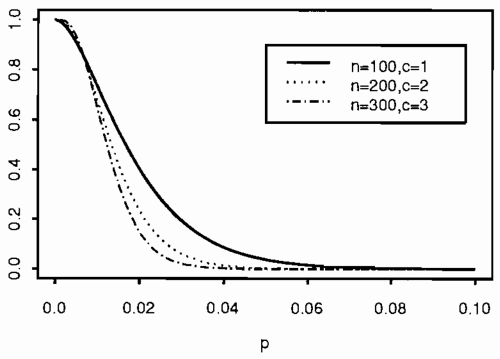
$$(3)\ P_a(p) = \begin{pmatrix} 300 \\ 0 \end{pmatrix} p^0(1-p)^{300} + \begin{pmatrix} 300 \\ 1 \end{pmatrix} p^1(1-p)^{299} + \begin{pmatrix} 300 \\ 2 \end{pmatrix} p^2(1-p)^{298} + \begin{pmatrix} 300 \\ 3 \end{pmatrix} p^3(1-p)^{297}$$
$$= (1-p)^{300} + 300p(1-p)^{299} + 44850p^2(1-p)^298 + 4455100p^3(1-p)^{297}$$

(b)

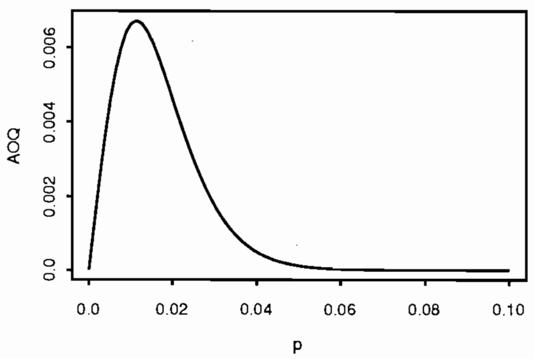
$$\begin{aligned} AOQ &= \left(1 - \frac{n}{N}\right) pP_a(p) \\ &= \left(1 - \frac{200}{10000}\right) p[(1-p)^{200} + 200p(1-p)^{199} + 19900p^2(1-p)^{198}] \\ &\qquad \text{if } n = 200, c = 2, N = 10000 \\ ATI(p) &= nP_a(p) + N(1 - P_a(p)) \\ &= (200 - 10000)[(1-p)^{200} + 200p(1-p)^{199} + 19900p^2(1-p)^{198}] + 10000 \end{aligned}$$

The AOQL is .00685.

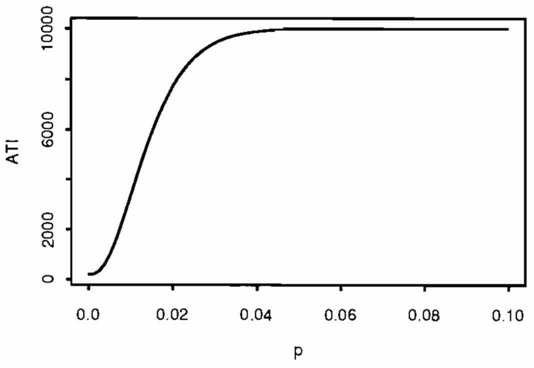
Approx OC curves from Perspective B



AOQ curve from perspective B, N=10000, n=200, c=2



ATI curve from perspective B, N=10000, n=200, c=2



6. (Problem 5.3 of the notes)

(a) We want to find an attributes single sampling plan such that $P_a = .95$ when $p = .01$ and $P_a = .10$ when $p = .03$.

$$\frac{Q_{2(C+1)}(1-.10)}{Q_{2(C+1)}(1-.95)} \approx \frac{.03}{.01} \qquad \Rightarrow \qquad \frac{Q_{2(C+1)}(.9)}{Q_{2(C+1)}(.05)} = 3$$

Noting that for $\nu = 16$, $Q_{16}(.9)/Q_{16}(.05) = 23.542/7.962 \approx 2.957$, we select $c = 7$. Therefore,

$$\begin{aligned} n &\approx Q_{16}(.9)/(2(.03)) = 23.542/.06 \approx 392 \\ n &\approx Q_{16}(.05)/(2(.01)) = 7.692/.02 \approx 385 \end{aligned}$$

so n is about 387.

(b)

$$\frac{Q_{2(C+1)}(1-.10)}{Q_{2(C+1)}(1-.95)} = \frac{3\times 10^{-6}}{1\times 10^{-6}} = 3$$

So, again, $c \approx 7$.

$$\begin{aligned} n &\approx Q_{16}(.05)/(2\times 10^{-6}) = 7.692/(2\times 10^{-6}) = 3846000 \\ n &\approx Q_{16}(.90)/(6\times 10^{-6}) = 23.542/(6\times 10^{-6}) = 3923667 \end{aligned}$$

so n is about 3,885,000.

7. (Problem 5.16 of the notes)

(a) $N = 3, k_1 = 1.5, k_2 = 10$
$X|p \sim Bin(n,p)$, $\quad p \sim G$ with pmf $g(p) = .5$ for $p = .1$ or $.2$, and $g(p) = 0$ otherwise.
Then the posterior of p given x is

$$f(p|x) = \frac{f(x|p)g(p)}{f(x|p=.1)g(.1) + f(x|p=.2)g(.2)} = \frac{p^x(1-p)^{n-x}}{(.1)^x(.9)^{n-x} + (.2)^x(.8)^{n-x}} \qquad p = .1, .2.$$

$$\begin{aligned} \text{If } n=1, x=0 \qquad f(p|x) &= \frac{1-p}{.9+.8} = \frac{1-p}{1.7} = \begin{cases} 9/17 & p = .1 \\ 8/17 & p = .2 \end{cases} \\ E(p|x) &= .1(9/17) + .2(8/17) = 25/170 = .147 \\ \text{If } n=1, x=1 \qquad f(p|x) &= \frac{p}{.1+.2} = \frac{p}{.3} = \begin{cases} 1/3 & p = .1 \\ 2/3 & p = .2 \end{cases} \\ E(p|x) &= .1(1/3) + .2(2/3) = 1/6 = .167 \\ \text{Note} \qquad c_G^{\text{opt}} &= \max\{x | E_G(p|X=x) \le k_1/k_2 = .15\} \end{aligned}$$

$$
\begin{aligned}
& \Rightarrow && c_G^{\text{opt}}(1) = 0 \qquad \text{since } E(p|x=0) < k_1/k_2 < E(p|x=1) \\
\text{If } n=2, x=0 \qquad & f(p|x) &=\;& \frac{(1-p)^2}{.81+.64} = \frac{(1-p)^2}{1.45} = \begin{cases} 81/145 & p=.1 \\ 64/145 & p=.2 \end{cases} \\
& E(p|x) &=\;& 81/1450 + 128/1450 = .1441 \\
\text{If } n=2, x=1 \qquad & f(p|x) &=\;& \frac{p(1-p)}{.09+.16} = \begin{cases} 9/25 & p=.1 \\ 16/25 & p=.2 \end{cases} \\
& E(p|x) &=\;& .164 \\
\text{If } n=2, x=2 \qquad & f(p|x) &=\;& \frac{p^2}{.01+.04} = \begin{cases} 1/5 & p=.1 \\ 4/5 & p=.2 \end{cases} \\
& E(p|x) &=\;& 9/50 = .18 \\
& \Rightarrow && c_G^{\text{opt}}(2) = 0 \text{ since only } x=0 \text{ yields } E(p|x) < .15 \\
\text{Similarly, if } n=3, & E(p|x=0) &=\;& \Pr(p=.1|x=0)(.1) + \Pr(p=.2|x=0)(.2) = .1413 \\
& E(p|x=1) &=\;& \Pr(p=.1|x=1)(.1) + \Pr(p=.2|x=1)(.2) = .161 \\
& \Rightarrow && c_G^{\text{opt}}(3) = 0 \text{ since } E(p|x=0) < .15 < E(p|x=1) \\
& && \text{and } E(p|x) \text{ is monotone nondecreasing}
\end{aligned}
$$

(b) With $N=3, c=0, k_1=1.5, k_2=10$,

$$
\begin{aligned}
E[\text{total cost}] &= k_1 N \left[1 + P_a(n,c,p)\left(1-\frac{n}{N}\right)\left(p\frac{k_2}{k_1}-1\right)\right], \\
&= 4.5\left[1 + P_a(n,0,p)\left(1-\frac{n}{3}\right)\left(p\frac{20}{3}-1\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
\text{If } n=0, c=0 \qquad E[\text{total cost}] &= \Pr(G=.1)E[TC(p=.1)] + \Pr(G=.2)E[TC(p=.2)] \\
&= (.5) \times E((\text{bad ones})(\text{cost/bad})) \\
&= (.5)(.1\times 3)10 + (.5)(.2\times 3)(10) \\
&= 4.5 \\
\text{If } n=1, c=0 \qquad E[\text{total cost}] &= 1.5 + .5[\Pr(X=1|p=.1)(\text{cost}|X=1) \\
&\quad + \Pr(X=0|p=.1)(\text{cost}|X=0) \\
&\quad + \Pr(X=1|p=.2)(\text{cost}|X=1) \\
&\quad + \Pr(X=0|p=.2)(\text{cost}|X=0)] \\
&= 4.45 \\
\text{If } n=2, c=0 \qquad E[\text{total cost}] &= \\
&= (1.5)(2) + .5[\Pr(X=1,2|p=.1)(\text{cost}|\text{reject}) \\
&\quad + \Pr(X=0|p=.1)(\text{cost}|\text{accept}) \\
&\quad + \Pr(X=1,2|p=.2)(\text{cost}|\text{reject}) \\
&\quad + \Pr(X=0|p=.2)(\text{cost}|\text{accept})] \\
&= 4.4575
\end{aligned}
$$

If $n = 3, c = 0$ $\quad E[\text{total cost}] = 3 \times 1.5 = 4.5$

Hence $c = 0, n = 1$ is the best plan.

8. (Problem 5.19 of the notes)
   $N = 5, k_1 = \$1, k_2 = \$100, p = .2.$

   (a) ALL : E(total cost) $= 5 \times \$1 = \$5$
   None : E(total cost) $= 5 \times .2 \times \$100 = \$100$
   Inspect until you find the defective, then quit.

$$\begin{aligned} E(\text{total cost}) &= \sum_{x=1}^{5} x \Pr(\text{defective found in } x\text{-th trial}) \\ &= 1 \cdot \frac{1}{5} + 2\frac{4}{5} \cdot \frac{1}{4} + 3 \cdot \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{1}{3} + 4 \cdot \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} + 5 \cdot \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} \cdot \frac{1}{1} \\ &= \frac{1}{5} \sum_{x=1}^{5} x = \$3 \end{aligned}$$

   (b) $p = .4, k_1 = \$1, k_2 = \$1000$
   ALL : E(total cost) $= 5 \times \$1 = \$5$
   None : E(total cost) $=$ Pr(all bad) $\times 1000 = (.4)^5 \times \$1000 = \$10.24$
   Sample only 1 item. If it is bad, replace it and then no penalty will be incurred since the lot has at least one good item.
   E(total cost) $= \$1$.

9. (Problem 5.27 of the notes)

   (a) The joint distribution of p and x:

	$x = 0$	$x = 1$	$g(p)$
$p = .0$	.3333	0	.333
$p = .1$	.2997	.0333	.333
$p = .2$	.2664	.0666	.333
$f(x)$	.8991	.0999	1.000

   The conditional distribution of p given x:

	$x = 0$	$x = 1$
$p = .0$	.3704	0
$p = .1$	.3333	.3333
$p = .2$	.2963	.6667

   (b)

   The cost of rejection is always 10 and the costs of acceptance for different x and p, and the expected costs for each given x are given in the tables below.
   For $n = 1$:

	$p = .0$	$p = .1$	$p = .2$	
$x = 0$	1	$1 + 10$	$1 + 20$	$1(.3704) + 11(.3333) + 21(.2963) = 10.26$
$x = 1$	NA	1	$1 + 10$	$1(.3333) + 11(.6667) = 7.67$

So the best plan that minimizes the expected cost rejects if $X = 0$ and accepts if $X = 1$.

For $n = 2$:

	$p = .0$	$p = .1$	$p = .2$	
$x = 0$	2	$2 + 10$	$2 + 20$	$2(.413) + 12(.330) + 22(.257) = 10.44$
$x = 1$	NA	2	$2 + 10$	$2(.360) + 12(.640) = 8.4$
$x = 2$	NA	NA	2	2

So the best plan that minimizes the expected cost rejects if $X = 0$ and accepts if $X = 1$ or 2.

(c) The lot contains very few defectives (0, 1, or 2 defectives with equal probability). For $n = 1$ or 2, if one or two items inspected are defective, then the remaining items are less likely to be defective than if the first one or two items inspected are not defective.

(d) For $n = 0$, the expected cost is 10 for either rejection or acceptance of the lot.

For $n = 1$, the expected cost for the best plan is $10(.8991) + 7.67(.0999) = 9.76$.

For $n = 2$, the expected cost for the best plan is $10(.807) + 8.4(.185) + 2(.007) = 9.64$.

Hence sample size $n = 2$ is the best here.