20170058 **Keonwoo Kim** · April 26, 2020
MAS555 Advanced Statistics

Homework 3

1. (Questions on Page 8 of Lecture 8 slide)

(i) Let Y_i be i.i.d. Bernoulli(p) variables. Then $\frac{1}{n}X_n \stackrel{d}{=} \bar{Y}_n$, the sample mean of Y_i's. Letting

$$\tilde{g}(x) = \log \frac{x + b/n}{1 - x + a/n}, \qquad \text{and}$$
$$g(x) = \log \frac{x}{1-x} + \left(\frac{b}{p} - \frac{a}{q}\right)\frac{1}{n},$$

we have

$$\begin{aligned}
\log \frac{p + \frac{b}{n}}{q + \frac{a}{n}} &= \log \frac{p + bh}{q + ah}\bigg|_{h=1/n} = \log \frac{p}{q} + \left(\frac{b}{p} - \frac{a}{q}\right) h + O(h^2)\bigg|_{h=1/n} \\
&= \log \frac{p}{q} + \left(\frac{b}{p} - \frac{a}{q}\right)\frac{1}{n} + O(n^{-2}), \\
\tilde{g}(x) - g(x) &= \log \frac{x + \frac{b}{n}}{1 - x + \frac{a}{n}} - \log \frac{x}{1-x} - \left(\frac{b}{p} - \frac{a}{q}\right)\frac{1}{n} \\
&= \log \frac{x + \frac{b}{n}}{1 - x + \frac{a}{n}} - \log \frac{p + \frac{b}{n}}{q + \frac{a}{n}} - \log \frac{x}{1-x} + \log \frac{p}{q} + O(n^{-2}).
\end{aligned}$$

Near $x = p$,

$$\begin{aligned}
\log \frac{x + \frac{b}{n}}{1 - x + \frac{a}{n}} - \log \frac{p + \frac{b}{n}}{q + \frac{a}{n}} &= \left(\frac{1}{p + \frac{b}{n}} + \frac{1}{q + \frac{a}{n}}\right)(x - p) \\
&\quad - \frac{1}{2}\left(\frac{1}{(p + \frac{b}{n})^2} - \frac{1}{(q + \frac{a}{n})^2}\right)(x-p)^2 + O((x-p)^3),
\end{aligned}$$

$$\begin{aligned}
&\log \frac{x}{1-x} - \log \frac{p}{q} \\
&\qquad = \left(\frac{1}{p} + \frac{1}{q}\right)(x - p) - \frac{1}{2}\left(\frac{1}{p^2} - \frac{1}{q^2}\right)(x-p)^2 + O((x-p)^3),
\end{aligned}$$

$$\therefore \log \frac{x + \frac{b}{n}}{1 - x + \frac{a}{n}} - \log \frac{p + \frac{b}{n}}{q + \frac{a}{n}} - \log \frac{x}{1-x} + \log \frac{p}{q}$$

$$
\begin{aligned}
&= \left(\frac{1}{p+\frac{b}{n}} - \frac{1}{p} + \frac{1}{q+\frac{a}{n}} - \frac{1}{q} \right)(x-p) \\
&\quad - \frac{1}{2}\left[\left(\frac{1}{(p+\frac{b}{n})^2} - \frac{1}{p^2} \right) - \left(\frac{1}{(q+\frac{a}{n})^2} - \frac{1}{q^2} \right) \right](x-p)^2 + O((x-p)^3) \\
&= -\frac{1}{n}\left(\frac{b}{p(p+\frac{b}{n})} + \frac{a}{q(q+\frac{a}{n})} \right)(x-p) \\
&\quad + \frac{1}{2}\cdot\frac{1}{n}\left(\frac{2bp+\frac{b^2}{n}}{p^2(p+\frac{b}{n})^2} - \frac{2aq+\frac{a^2}{n}}{q^2(q+\frac{a}{n})^2} \right)(x-p)^2 + O((x-p)^3) \\
&= O(n^{-1})(x-p) + O(n^{-1})(x-p)^2 + O((x-p)^3)
\end{aligned}
$$

so that

$$
\begin{aligned}
\tilde{g}(x) - g(x) &= \log\frac{x+\frac{b}{n}}{1-x+\frac{a}{n}} - \log\frac{p+\frac{b}{n}}{q+\frac{a}{n}} - \log\frac{x}{1-x} + \log\frac{p}{q} + O(n^{-2}) \\
&= O(n^{-1})(x-p) + O(n^{-1})(x-p)^2 + O((x-p)^3) + O(n^{-2}).
\end{aligned}
$$

Note the CLT implies that $\sqrt{n}(\bar{Y}_n - p)$ has $\mathcal{N}(0, pq)$ as the asymptotic distribution, which implies $\bar{Y}_n - p = O_p(n^{-1/2})$. Thus,

$$
\begin{aligned}
&\tilde{g}(\bar{Y}_n) - g(\bar{Y}_n) \\
&= O(n^{-1})O_p(n^{-1/2}) + O(n^{-1})O_p(n^{-1}) + O_p(n^{-3/2}) + O(n^{-2}) \\
&= O_p(n^{-3/2}),
\end{aligned}
$$

which is $o_p(1/n)$. Now we have

$$
\begin{aligned}
g(\bar{Y}_n) = \left[\log\frac{p}{q} + \left(\frac{b}{p} - \frac{a}{q}\right)\frac{1}{n}\right] &+ \frac{\sigma}{\sqrt{n}}\left(\frac{1}{p} + \frac{1}{q}\right) Z_n \\
&+ \frac{\sigma^2}{2n}\left(-\frac{1}{p^2} + \frac{1}{q^2}\right) Z_n^2 + o_p(1),
\end{aligned}
$$

where $\sigma^2 = pq$ and $Z_n = \sqrt{n}(\bar{Y}_n - p)/\sigma$. Thus, using $\tilde{g}(\frac{1}{n}X_n) - g(\frac{1}{n}X_n) \stackrel{d}{=} \tilde{g}(\bar{Y}_n) - g(\bar{Y}_n) = o_p(1)$,

$$
\begin{aligned}
\log\frac{X_n+b}{n-X_n+a} &= \left[\log\frac{p}{q} + \left(\frac{b}{p} - \frac{a}{q}\right)\frac{1}{n}\right] \\
&\qquad + \frac{\sigma}{\sqrt{n}}\left(\frac{1}{p} + \frac{1}{q}\right)\tilde{Z}_n + \frac{\sigma^2}{2n}\left(-\frac{1}{p^2} + \frac{1}{q^2}\right)\tilde{Z}_n^2 + o_p(1)
\end{aligned}
$$

where $\sigma^2 = pq$ and $\tilde{Z}_n \coloneqq \sqrt{n}(\frac{1}{n}X_n - p)/\sigma$. Hence,

$$
\begin{aligned}
W_n &= \left[\log\frac{p}{q} + \left(\frac{b}{p} - \frac{a}{q}\right)\frac{1}{n}\right] + \frac{\sigma}{\sqrt{n}}\left(\frac{1}{p} + \frac{1}{q}\right)\tilde{Z}_n + \frac{\sigma^2}{2n}\left(-\frac{1}{p^2} + \frac{1}{q^2}\right)\tilde{Z}_n^2 \\
&= \left[\log\frac{p}{q} + \left(\frac{b}{p} - \frac{a}{q}\right)\frac{1}{n}\right] \\
&\qquad + \left(\frac{1}{p} + \frac{1}{q}\right)\left(\frac{1}{n}X_n - p\right) + \frac{1}{2}\left(-\frac{1}{p^2} + \frac{1}{q^2}\right)\left(\frac{1}{n}X_n - p\right)^2
\end{aligned}
$$

could be an answer.

(ii) Using the limit theorem (7–i) in the lecture note to $g(\bar{Y}_n) \stackrel{d}{=} g(\frac{1}{n}X_n) = W_n + o_p(n^{-1})$,

$$
\begin{aligned}
\mathrm{E}W_n &= g(p) + \frac{\sigma^2}{2n}g''(p) \\
&= \log\frac{p}{q} + \left(\frac{b}{p} - \frac{a}{q}\right)\frac{1}{n} + \frac{pq}{2n}\left(-\frac{1}{p^2} + \frac{1}{q^2}\right) \\
&= \log\frac{p}{q} + \left(\frac{b - \frac{1}{2}}{p} + \frac{\frac{1}{2} - a}{q}\right)\frac{1}{n}, \\
\operatorname{Var} W_n &= \frac{1}{n}(g'(p))^2\sigma^2 + o_p(n^{-1}) \\
&= \frac{pq}{n}\left(\frac{1}{p} + \frac{1}{q}\right)^2 + o_p(n^{-1}) \\
&= \frac{pq}{n}\left(\frac{p+q}{pq}\right)^2 + o_p(n^{-1}) \\
&= \frac{1}{n}\cdot\frac{1}{pq} + o_p(n^{-1}). && (\because p + q = 1)
\end{aligned}
$$

2. As in the lecture material, with the initial unbiased estimator $T_0 = \mathbf{1}_{(X_1 \le u)}$, the Rao–Blackwellization of T_0 given the CSS $(\bar{X}, \sum_i (X_i - \bar{X})^2)'$ is:

$$
\begin{aligned}
&T_1\Big(\bar{x}, \sum_i (x_i - \bar{x})^2\Big) \\
&= P\Big(X_1 \le u \,\Big|\, \bar{X} = \bar{x}, \sum (X_i - \bar{X})^2 = \sum (x_i - \bar{x})^2\Big) \\
&= P\left(\frac{X_1 - \bar{X}}{\sqrt{\sum (X_i - \bar{X})^2}} \le \frac{u - \bar{x}}{\sqrt{\sum (x_i - \bar{x})^2}} \,\middle|\, \bar{X} = \bar{x}, \sum (X_i - \bar{X})^2 = \sum (x_i - \bar{x})^2\right) \\
&= P\left(\frac{X_1 - \bar{X}}{\sqrt{\sum (X_i - \bar{X})^2}} \le \frac{u - \bar{x}}{\sqrt{\sum (x_i - \bar{x})^2}}\right). && \text{(Basu's theorem)}
\end{aligned}
$$

First of all, since $(X_1-\bar{X})\big(\sum(X_i-\bar{X})^2\big)^{-1/2}$ does not depend on the values of μ and σ^2, without loss of generality, we may assume $\mu=0$ and $\sigma^2=1$. Here, $X_1-\bar{X}$ follows a normal distribution, where the mean is $\mathrm{E}(X_1-\bar{X})=0$ and the variance is

$$\begin{aligned}\mathrm{Var}(X_1-\bar{X}) &= \mathrm{Var}\left(\frac{n-1}{n}X_1-\frac{1}{n}\sum_{j=2}^n X_j\right)\\ &= \left(\frac{n-1}{n}\right)^2+\sum_{j=2}^n\frac{1}{n^2}=\frac{n-1}{n}.\end{aligned}$$

Therefore, $\sqrt{\frac{n}{n-1}}(X_1-\bar{X})\sim\mathcal{N}(0,1)$, and hence $\Xi_1\coloneqq\frac{n}{n-1}(X_1-\bar{X})^2\sim\chi_1^2$. Since

$$\Xi_1+\Xi_2\coloneqq\sum_{j=1}^n(X_j-\bar{X})^2\sim\chi_{n-1}^2,$$

we have

$$\frac{\frac{n}{n-1}(X_1-\bar{X})^2}{\sum_{j=1}^n(X_j-\bar{X})^2}=\frac{\Xi_1}{\Xi_1+\Xi_2}\sim\mathrm{Beta}\left(\frac{1}{2},\frac{n-2}{2}\right)$$

since $\Xi_2=-\frac{1}{n-1}(X_1-\bar{X})^2+\sum_{j=2}^n(X_j-\bar{X})^2\sim\chi_{n-2}^2$.

Note that $-X_1,\dots,-X_n$ has the same property as $X_1,\dots,X_n$, yielding

$$\frac{X_1-\bar{X}}{\sqrt{\sum(X_i-\bar{X})^2}}\overset{d}{=}\frac{(-X_1)-(-\bar{X})}{\sqrt{\sum((-X_i)-(-\bar{X}))^2}}=-\frac{X_1-\bar{X}}{\sqrt{\sum(X_i-\bar{X})^2}}.$$

Hence,

$$\begin{aligned}&P\left(\left(\frac{X_1-\bar{X}}{\sqrt{\sum(X_i-\bar{X})^2}}\right)^2>\left(\frac{u-\bar{x}}{\sqrt{\sum(x_i-\bar{x})^2}}\right)^2\right)\\ &=P\left(\frac{X_1-\bar{X}}{\sqrt{\sum(X_i-\bar{X})^2}}>\frac{|u-\bar{x}|}{\sqrt{\sum(x_i-\bar{x})^2}}\right)+P\left(\frac{X_1-\bar{X}}{\sqrt{\sum(X_i-\bar{X})^2}}<-\frac{|u-\bar{x}|}{\sqrt{\sum(x_i-\bar{x})^2}}\right)\\ &=P\left(\frac{X_1-\bar{X}}{\sqrt{\sum(X_i-\bar{X})^2}}>\frac{|u-\bar{x}|}{\sqrt{\sum(x_i-\bar{x})^2}}\right)+P\left(-\frac{X_1-\bar{X}}{\sqrt{\sum(X_i-\bar{X})^2}}>\frac{|u-\bar{x}|}{\sqrt{\sum(x_i-\bar{x})^2}}\right)\\ &=2P\left(\frac{X_1-\bar{X}}{\sqrt{\sum(X_i-\bar{X})^2}}>\frac{|u-\bar{x}|}{\sqrt{\sum(x_i-\bar{x})^2}}\right)\end{aligned}$$

$$= 2P\left(\frac{X_1 - \bar{X}}{\sqrt{\sum(X_i - \bar{X})^2}} < -\frac{|u - \bar{x}|}{\sqrt{\sum(x_i - \bar{x})^2}}\right).$$

The probability in the first line of the equation above can be evaluated as follows:

$$\begin{aligned}
&P\left(\left(\frac{X_1 - \bar{X}}{\sqrt{\sum(X_i - \bar{X})^2}}\right)^2 > \left(\frac{u - \bar{x}}{\sqrt{\sum(x_i - \bar{x})^2}}\right)^2\right) \\
&\quad = P\left(\frac{\frac{n}{n-1}(X_1 - \bar{X})^2}{\sum(X_i - \bar{X})^2} > \frac{\frac{n}{n-1}(u - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right) \\
&\quad = 1 - \mathrm{cdf}_{\mathrm{Beta}(\frac{1}{2}, \frac{n-2}{2})}\left(\frac{\frac{n}{n-1}(u - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right),
\end{aligned}$$

When $u \geq \bar{x}$, we have

$$\begin{aligned}
&P\left(\frac{X_1 - \bar{X}}{\sqrt{\sum(X_i - \bar{X})^2}} \leq \frac{u - \bar{x}}{\sqrt{\sum(x_i - \bar{x})^2}}\right) \\
&= 1 - P\left(\frac{X_1 - \bar{X}}{\sqrt{\sum(X_i - \bar{X})^2}} > \frac{|u - \bar{x}|}{\sqrt{\sum(x_i - \bar{x})^2}}\right) \\
&= \frac{1}{2} + \frac{1}{2}\,\mathrm{cdf}_{\mathrm{Beta}(\frac{1}{2}, \frac{n-2}{2})}\left(\frac{\frac{n}{n-1}(u - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right).
\end{aligned}$$

When $u \leq \bar{x}$, we have

$$\begin{aligned}
&P\left(\frac{X_1 - \bar{X}}{\sqrt{\sum(X_i - \bar{X})^2}} \leq \frac{u - \bar{x}}{\sqrt{\sum(x_i - \bar{x})^2}}\right) \\
&= P\left(\frac{X_1 - \bar{X}}{\sqrt{\sum(X_i - \bar{X})^2}} < -\frac{|u - \bar{x}|}{\sqrt{\sum(x_i - \bar{x})^2}}\right) \\
&= \frac{1}{2} - \frac{1}{2}\,\mathrm{cdf}_{\mathrm{Beta}(\frac{1}{2}, \frac{n-2}{2})}\left(\frac{\frac{n}{n-1}(u - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right).
\end{aligned}$$

Therefore,

$$T_1 = \frac{1}{2} + \frac{1}{2}\,\mathrm{sgn}(u - \bar{X}) \cdot \mathrm{cdf}_{\mathrm{Beta}(\frac{1}{2}, \frac{n-2}{2})}\left(\frac{\frac{n}{n-1}(u - \bar{X})^2}{\sum(X_i - \bar{X})^2}\right)$$

and T_1 is the unique UMVUE of $P(X_1 \leq u)$ due to Lehmann–Scheffé theorem.

6.17. Note that

$$P_\theta(X_1 = x_1, \ldots, X_n = x_n) = \prod_{j=1}^{n} \theta(1-\theta)^{x_j - 1} = \theta^n (1-\theta)^{\sum_j x_j - n}.$$

Therefore, by the factorization theorem, $\sum X_i$ is sufficient for θ.

$\sum X_i$ follows the following distribution:

$$P\left(\sum_{i=1}^{n} X_i = k\right) = \binom{k-1}{k-n} \theta^n (1-\theta)^{k-n} \qquad (k = n,\, n+1,\, \cdots).$$

When $n = 1$, the equation holds with $P(X_1 = k) = \theta(1-\theta)^{k-1}$. The induction step is proceeded as follows:

$$\begin{aligned} P\left(\sum_{i=1}^{n} X_i = k\right) &= \sum_{r=1}^{k-n+1} P\left(\sum_{i=1}^{n-1} X_i = k - r,\ X_n = r\right) \\ &= \sum_{r=1}^{k-n+1} \binom{k-r-1}{k-r-n+1} \theta^{n-1}(1-\theta)^{k-r-n+1}\, \theta(1-\theta)^{r-1} \\ &= \theta^n (1-\theta)^{k-n} \sum_{r=1}^{k-n+1} \binom{k-r-1}{k-r-n+1} \\ &= \theta^n (1-\theta)^{k-n} \binom{k-1}{k-n}. \end{aligned}$$

Now, denote $T = \sum_{i=1}^{n} X_i$. Assume for any $0 < \theta < 1$,

$$\mathrm{E}_\theta\left[g(T)\right] = \sum_{k=n}^{\infty} g(k) \binom{k-1}{k-n} \theta^n (1-\theta)^{k-n} = 0.$$

This implies $g(k)\binom{k-1}{k-n}(1-\theta)^k = 0$ for any $k \geq n$, i.e., $g(k) = 0$ for any $k \geq n$. Thus, $T = \sum X_i$ is complete.

6.30. (a) The sufficiency is proven using the factorization theorem:

$$f(x_1, \ldots, x_n | \mu) = \prod_{j=1}^{n} e^{-(x_j - \mu)} \mathbf{1}_{(\mu < x_j)} = e^{n\mu} \mathbf{1}_{(\mu < x_{(1)})} \cdot \exp\Bigl(-\sum_{j=1}^{n} x_j\Bigr)$$

where $x_{(1)} = \min_{1 \leq j \leq n} x_j$.

To see the completeness, assume for any μ,

$$\mathrm{E}g(X_{(1)}|\mu) = \int_{\mu}^{\infty} g(x)\, ne^{-n(x-\mu)}\, dx = 0.$$

Then we have $g(x)\, ne^{-n(x-\mu)} = 0$ a.e., that is, $g(x) = 0$ a.e. Thus $g(X_{(1)}) = 0$ a.s., for any μ. This shows the completeness.

(b) Since the common distribution of X_i's is in a location parameter family, S^2 is an ancillary statistic, as $X_j - \bar{X} = (X_j + a) - (\bar{X} + a)$ is independent of μ. Therefore, since $X_{(1)}$ is a CSS and S^2 is ancillary, they are independent due to Basu's theorem.

7·37· (a)

$$\begin{aligned}
d_{\mathrm{P}}^{r}(cx_1, \ldots, cx_n) &= \frac{\int_0^\infty t^{n+r-1} \prod_{i=1}^n f(tcx_i)\, dt}{\int_0^\infty t^{n+2r-1} \prod_{i=1}^n f(tcx_i)\, dt} \\
&= c^r \frac{\int_0^\infty (ct)^{n+r-1} \prod_{i=1}^n f(ct \cdot x_i)\, c\, dt}{\int_0^\infty (ct)^{n+2r-1} \prod_{i=1}^n f(ct \cdot x_i)\, c\, dt} \\
&= c^r \frac{\int_0^\infty t^{n+r-1} \prod_{i=1}^n f(tx_i)\, dt}{\int_0^\infty t^{n+2r-1} \prod_{i=1}^n f(tx_i)\, dt} \\
&= c^r d_{\mathrm{P}}^{r}(x_1, \ldots, x_n).
\end{aligned}$$

(b) In this case, $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2)$ with $r = 2$. Since $\int_0^\infty t^{z-1} e^{-st}\, dt = s^{-z}\Gamma(z)$, we have

$$\begin{aligned}
d_{\mathrm{P}}^{2}(x_1, \ldots, x_n) &= \frac{\int_0^\infty t^{n+1} (2\pi)^{-n/2} \exp\bigl(-t^2 \sum_{i=1}^n x_i^2\bigr)\, dt}{\int_0^\infty t^{n+3} (2\pi)^{-n/2} \exp\bigl(-t^2 \sum_{i=1}^n x_i^2\bigr)\, dt} \\
&= \frac{\int_0^\infty t^{n+1} \exp\bigl(-t^2 \sum_{i=1}^n x_i^2\bigr)\, dt}{\int_0^\infty t^{n+3} \exp\bigl(-t^2 \sum_{i=1}^n x_i^2\bigr)\, dt} \\
&= \frac{\int_0^\infty u^{n/2} \exp\bigl(-u \sum_{i=1}^n x_i^2\bigr)\, du}{\int_0^\infty u^{(n/2)+1} \exp\bigl(-u \sum_{i=1}^n x_i^2\bigr)\, du} \\
&= \frac{\bigl(\sum_{i=1}^n x_i^2\bigr)^{-(n/2)-1} \Gamma((n/2)+1)}{\bigl(\sum_{i=1}^n x_i^2\bigr)^{-(n/2)-2} \Gamma((n/2)+2)} \\
&= \frac{\sum_{i=1}^n x_i^2}{(n/2)+1} = \frac{2}{n+2} \sum_{i=1}^n x_i^2.
\end{aligned}$$

(c) In this case, $f(x) = \exp(-x)$ with $\sigma = \beta$ and $r = 1$.

$$
\begin{aligned}
d_{\mathrm{P}}^1(x_1, \ldots, x_n) &= \frac{\int_0^\infty t^n \prod_{i=1}^n \exp(-tx_i)\, dt}{\int_0^\infty t^{n+1} \prod_{i=1}^n \exp(-tx_i)\, dt} \\
&= \frac{\int_0^\infty t^n \exp(-t \sum x_i)\, dt}{\int_0^\infty t^{n+1} \exp(-t \sum x_i)\, dt} \\
&= \frac{(\sum x_i)^{-n-1} \Gamma(n+1)}{(\sum x_i)^{-n-2} \Gamma(n+2)} \\
&= \frac{1}{n+1} \sum_i x_i.
\end{aligned}
$$

(d) In this case, $f(x) = \mathbf{1}_{(0<x<1)}$ with $\sigma = \theta$ and $r = 1$.

$$
\begin{aligned}
d_{\mathrm{P}}^1(x_1, \ldots, x_n) &= \frac{\int_0^\infty t^n \prod_{i=1}^n \mathbf{1}_{(0<tx_i<1)}\, dt}{\int_0^\infty t^{n+1} \prod_{i=1}^n \mathbf{1}_{(0<tx_i<1)}\, dt} \\
&= \frac{\int_0^{1/\max_i x_i} t^n\, dt}{\int_0^{1/\max_i x_i} t^{n+1}\, dt} \\
&= \frac{n+2}{n+1} \max_i x_i.
\end{aligned}
$$

7.50. (a) $\mathrm{E}[a\bar{X} + (1-a)cS] = a\mathrm{E}\bar{X} + (1-a)\mathrm{E}(cS) = a\theta + (1-a)\theta = \theta$.

(b) Note that the sample mean $\bar{X}$ and the sample variance S^2 are independent provided the normality of the distribution. Thus, so are $\bar{X}$ and S. Thus,

$$
\mathrm{Var}(a\bar{X} + (1-a)cS) = a^2 \mathrm{Var}(\bar{X}) + (1-a)^2 c^2 \mathrm{Var}(S).
$$

Note that $(n-1)S^2/\theta^2 \sim \chi^2_{n-1}$, so that the mean of $\sqrt{(n-1)S^2/\theta^2}$ is as follows:

$$
\begin{aligned}
\mathrm{E}\left[\sqrt{\frac{(n-1)S^2}{\theta^2}}\right] &= \int_0^\infty \sqrt{x}\, \frac{x^{(n-1)/2-1} e^{-x/2}}{2^{(n-1)/2}\Gamma((n-1)/2)}\, dx \\
&= \int_0^\infty \frac{x^{n/2-1} e^{-x/2}}{2^{n/2}\Gamma(n/2)} \frac{2^{n/2}\Gamma(n/2)}{2^{(n-1)/2}\Gamma((n-1)/2)}\, dx \\
&= \sqrt{2} \frac{\Gamma(n/2)}{\Gamma((n-1)/2)}.
\end{aligned}
$$

Therefore,

$$\begin{aligned}
\mathrm{Var}[S] &= \mathrm{E}[S^2] - (\mathrm{E}S)^2 \\
&= \theta^2 - \left(\sqrt{\frac{2}{n-1}}\,\theta\,\frac{\Gamma(n/2)}{\Gamma((n-1)/2)}\right)^2 \\
&= \theta^2 - (\theta/c)^2 . \\
\therefore \mathrm{Var}(a\bar{X} + (1-a)cS) &= a^2\,\frac{\theta^2}{n} + (1-a)^2(c^2-1)\theta^2.
\end{aligned}$$

At the value of a minimizing the variance of $a\bar{X} + (1-a)cS$, we have $2a\,\frac{\theta^2}{n} - 2(1-a)(c^2-1)\theta^2 = 0$, i.e.,

$$a = \frac{c^2-1}{\frac{1}{n} + c^2 - 1}.$$

(c) As the joint pdf of X_i's is

$$\begin{aligned}
f_\theta(x_1, \ldots, x_n) &= (2\pi)^{-n/2}\theta^{-n}\exp\left[-\frac{1}{2\theta^2}\sum_i (x_i-\theta)^2\right] \\
&= (2\pi)^{-n/2}\theta^{-n}\exp\left[-\frac{1}{2\theta^2}\sum x_i^2 + \frac{1}{\theta}\sum x_i - \frac{n}{2}\right] \\
&= (2\pi)^{-n/2}\theta^{-n}\exp\left[-\frac{1}{2\theta^2}((n-1)s^2 + n\bar{x}^2) + \frac{1}{\theta}n\bar{x} - \frac{n}{2}\right]
\end{aligned}$$

where $\bar{x} = \sum x_i$ and $s^2 = (n-1)^{-1}\sum(x_i-\bar{x})^2 = (n-1)^{-1}(\sum x_i^2 - n\bar{x}^2)$, $\bar{X}, S^2$ is sufficient for θ. However, for

$$g(\bar{x}, s^2) = \frac{1}{n}\,s^2 - \frac{1}{n+1}(\bar{x})^2 \not\equiv 0 \quad \text{(in an almost sure sense)},$$

we have

$$\mathrm{E}_\theta[g(\bar{X}, S^2)] = \frac{1}{n}\,\mathrm{E}S^2 - \frac{1}{n+1}\mathrm{E}(\bar{X}^2) = \frac{1}{n}\theta^2 - \frac{1}{n+1}\left(\frac{\theta^2}{n} + \theta^2\right) = 0.$$

This proves that $(\bar{X}, S^2)$ is not a complete sufficient statistic for θ.