Stat 5102 (Geyer) Midterm 1

Problem 1

The basic fact this problem uses is

$$\frac{\overline{X}_n - \mu}{S_n/\sqrt{n}} \sim t(n-1)$$

(Corollary 7.25 in the notes), which, since $\mu = 0$, specializes to

$$\frac{\overline{X}_n}{S_n/\sqrt{n}} \sim t(n-1) \tag{1}$$

in this problem. This is the only exact result we have involving both $\overline{X}_n$ and S_n, so nothing else is of any use.

To use (1), we must put the event of interest $\overline{X}_n < S_n$ in a form related to the left hand side of (1). Clearly, this is equivalent to

$$\frac{\overline{X}_n}{S_n/\sqrt{n}} < \sqrt{n} = 3$$

So the probability we need to find $P(Y < 3)$ where Y is a $t(n-1)$ random variable ($n - 1 = 8$ degrees of freedom).

From Table IIIa in Lindgren $P(Y > 3) = 0.009$, so $P(Y < 3) = 1 - 0.009 = 0.991$.

Problem 2

To use the method of moments, we first need to find some moments. Since this is not a "brand name" distribution, we must integrate to find the moments. The obvious moment to try first is the first moment (the mean)

$$\begin{aligned}
\mu &= \int_0^1 x f_\beta(x)\,dx \\
&= \frac{2}{1+\beta}\int_0^1 [x + (\beta - 1)x^2]\,dx \\
&= \frac{2}{1+\beta}\left[\frac{x^2}{2} + (\beta - 1)\frac{x^3}{3}\right]_0^1 \\
&= \frac{1+2\beta}{3+3\beta}
\end{aligned}$$

Solving for β as a function of μ, we get

$$\beta = \frac{3\mu - 1}{2 - 3\mu}$$

(the numerator and denominator are both positive because $1/3 < \mu < 2/3$.)

Either way, we get a method of moments estimator by plugging in $\overline{X}_n$ for μ

$$\hat{\beta}_n = \frac{3\overline{X}_n - 1}{2 - 3\overline{X}_n}$$

Problem 3

(a) The asymptotic distribution of $\overline{X}_n$ is, as usual, by the CLT,

$$\overline{X}_n \approx \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

Plugging in $\sigma^2 = \mu$ gives

$$\overline{X}_n \approx \mathcal{N}\left(\mu, \frac{\mu}{n}\right).$$

(b) The asymptotic distribution of V_n is, by Corollary 7.17 in the notes,

$$V_n \approx \mathcal{N}\left(\mu_2, \frac{\mu_4 - \mu_2^2}{n}\right).$$

where $\mu_2 = \sigma^2 = \mu$ and $\mu_4 = \mu + 3\mu^2$ are given in the problem statement. Plugging these in gives

$$V_n \approx \mathcal{N}\left(\mu, \frac{\mu + 2\mu^2}{n}\right).$$

(c) The ARE is the ratio of the asymptotic variances, either $1 + 2\mu$ or the the reciprocal $1/(1 + 2\mu)$, depending on which way you write it.

(d) The better estimator is the one with the smaller asymptotic variance, in this case $\overline{X}_n$.

Problem 4

This is a problem for the delta method. We know from the properties of the exponential distribution

$$E(X_i) = \frac{1}{\lambda}$$

and

$$\operatorname{var}(X_i) = \frac{1}{\lambda^2}$$

Hence the CLT says in this case

$$\overline{X}_n \approx \mathcal{N}\left(\frac{1}{\lambda}, \frac{1}{n\lambda^2}\right)$$

For any differentiable function g, the delta method says

$$g(\overline{X}_n) \approx \mathcal{N}\left(g\left(\frac{1}{\lambda}\right), g'\left(\frac{1}{\lambda}\right)^2 \frac{1}{n\lambda^2}\right)$$

The g such that $W_n = g(\overline{X}_n)$ is

$$g(x) = \frac{x}{1 + x} \tag{2}$$

which has derivative

$$g'(x) = \frac{1}{(1 + x)^2} \tag{3}$$

so

$$g\left(\frac{1}{\lambda}\right) = \frac{1}{1 + \lambda}$$

and

$$g'\left(\frac{1}{\lambda}\right) = \frac{\lambda^2}{(1 + \lambda)^2}$$

and

$$W_n \approx \mathcal{N}\left(\frac{1}{1 + \lambda}, \frac{\lambda^2}{n(1 + \lambda)^4}\right)$$

Problem 5

There are several different ways to proceed here.

Using a Confidence Interval for the Mean

The mean is $\mu = 1/p$. Thus we could just get a confidence interval for μ and take reciprocals of the endpoints to get a confidence interval for p.

A confidence interval for μ can be found using Theorem 9.8 in the notes. From Section B.1.8 of the notes

$$\operatorname{var}(X_i) = \frac{1 - p}{p^2} = \mu(\mu - 1)$$

so by the LLN and the continuous mapping theorem

$$S_n = \sqrt{\overline{X}_n(\overline{X}_n - 1)}$$

is a consistent estimator of the population standard deviation σ needed for the theorem. The theorem gives

$$\overline{X}_n \pm 1.96\sqrt{\frac{\overline{X}_n(\overline{X}_n - 1)}{n}}$$

as an asymptotic 95% confidence interval for μ. So

$$\frac{1}{\overline{X}_n + 1.96\sqrt{\frac{\overline{X}_n(\overline{X}_n - 1)}{n}}} < p < \frac{1}{\overline{X}_n - 1.96\sqrt{\frac{\overline{X}_n(\overline{X}_n - 1)}{n}}}$$

is an asymptotic 95% confidence interval for $p = 1/\mu$. Plugging in the numbers gives

$$0.180 < p < 0.256$$

Using the Delta Method

The obvious point estimator for p is

$$\hat{p}_n = \frac{1}{\overline{X}_n}$$

The CLT says

$$\overline{X}_n \approx \mathcal{N}\left(\frac{1}{p}, \frac{(1-p)}{np^2}\right)$$

Applying the delta method with the transformation

$$g(u) = \frac{1}{u}$$

with derivative

$$g'(u) = -\frac{1}{u^2}$$

gives

$$g\left(\frac{1}{p}\right) = p$$

and

$$g'\left(\frac{1}{p}\right) = -p^2$$

and

$$\hat{p}_n \approx \mathcal{N}\left(p, \frac{p^2(1-p)}{n}\right)$$

which gives an asymptotic 95% confidence interval

$$\hat{p}_n \pm 1.96\sqrt{\frac{\hat{p}_n^2(1-\hat{p}_n)}{n}}$$

Plugging in the numbers gives

$$0.2114 \pm 0.03680$$

or

$$0.1746 < p < 0.2482$$

which is pretty close to the other interval.

Solving Quadratic Inequalities

The really hard way to do this problem is to start with

$$\overline{X}_n \approx \mathcal{N}\left(\frac{1}{p}, \frac{1-p}{np^2}\right)$$

and standardize giving the asymptotically standard normal quantity

$$\frac{\overline{X}_n - \frac{1}{p}}{\sqrt{\frac{1-p}{np^2}}}$$

from which we conclude that the set of p such that

$$\left| \frac{\overline{X}_n - \frac{1}{p}}{\sqrt{\frac{1-p}{np^2}}} \right| < 1.96$$

is an asymptotic 95% confidence interval for p. It turns out this is solvable, equivalent to

$$\begin{aligned} 1.96^2 &> \left| \frac{\overline{X}_n - \frac{1}{p}}{\sqrt{\frac{1-p}{np^2}}} \right|^2 \\ &= \frac{n(\overline{X}_n p - 1)^2}{1-p} \end{aligned}$$

Or, writing $z = 1.96$, we see the confidence interval has endpoints satisfying the quadratic equation

$$(1-p)z^2 = n(\overline{X}_n p - 1)^2$$

which has roots

$$\frac{2n\overline{X}_n - z^2 \pm z\sqrt{4n\overline{X}_n + z^2 - 4n\overline{X}_n}}{2n\overline{X}_n^2}$$

or

$$0.17375 < p < 0.247366$$