ANSWER TO OCTOBER 12 EXERCISE

CHRIS SIMS AND MATT WEINBERG

$\overline{Y} = capacity$
T i.i.d. observations on output Y_t.

$$\log Y_t = \log \overline{Y} - \varepsilon_t$$

where $\left\{\varepsilon_t | \overline{Y}\right\} \sim \Gamma(1,1)$ at each t.

(a) Find an unbiased estimator of $\log \overline{Y}$.
$E[\ln Y_t | \overline{Y}] = \ln \overline{Y} - E[\varepsilon_t | \overline{Y}] = \ln \overline{Y} - 1$. *Therefore an unbiased estimator is given by*

$$\widehat{\log \overline{Y}} = \frac{1}{T} \sum_{t=1}^{T} \log Y_t + 1$$

because

$$\begin{aligned} E[\widehat{\log \overline{Y}}] &= \frac{1}{T} \sum_{t=1}^{T} E[\log Y_t] + 1 \\ &= \log \overline{Y} + \frac{1}{T} \sum_{t=1}^{T} E[\varepsilon_t] + 1 \\ &= \log \overline{Y} + \frac{1}{T} \sum_{t=1}^{T} E\left[E[\varepsilon_t | \overline{Y}]\right] + 1 \\ &= \log \overline{Y} \end{aligned}$$

(b) Find an unbiased estimator of $\overline{Y}$.

$$E[Y_t] = E[\overline{Y} e^{-\varepsilon_t}] = E\left[E[\overline{Y} e^{-\varepsilon_t} | \overline{Y}]\right] = E\left[\overline{Y} E[e^{-\varepsilon_t} | \overline{Y}]\right]$$

$$E[e^{-\varepsilon_t} | \overline{Y}] = \int_0^\infty e^{-\varepsilon_t} e^{-\varepsilon_t} d\varepsilon_t = \int_0^\infty e^{-2\varepsilon_t} d\varepsilon_t = -\frac{1}{2} e^{-2\varepsilon_t} |_{\varepsilon=0}^{\infty} = \frac{1}{2}$$

So $2E[Y_t] = E[\overline{Y}]$*, so take our estimator as*

$$\widehat{\overline{Y}} = \frac{2}{T} \sum_{t=1}^{T} e^{\log Y_t}$$

To verify that it is not biased, note that

$$\begin{aligned} E[\widehat{\overline{Y}}] &= E[\frac{2}{T}\sum_{t=1}^{T}\overline{Y}e^{-\varepsilon_t}] \\ &= E[\frac{2}{T}\sum_{t=1}^{T}\overline{Y}E\left(e^{-\varepsilon_t}|\overline{Y}\right)] \\ &= E\left\{\frac{2}{T}\frac{T\overline{Y}}{2}\right\} = \overline{Y} \end{aligned}$$

(c) Find a sufficient statistic for $\overline{Y}$. *One can check using the change of variable formula that* Y_t *has a* $Uniform\left(0,\overline{Y}\right)$ *distribution. then*

$$f\left(Y_t|\overline{Y}\right) = \frac{1}{\overline{Y}}$$

and the likelihood is

$$L\left(\{Y_t\}_{t=1}^{T}|\overline{Y}\right) = \prod_{t=1}^{T}\left(\frac{1}{\overline{Y}}\right)^{T}\mathbf{1}_{\max(Y_t)<\overline{Y}}$$

(d) $\overline{Y}_{MLE} = \max\{Y_1,...,Y_T\}$.

(e) With a uniform prior,

$$p\left(Y|\overline{Y}\right) \propto \left(\frac{1}{\overline{Y}}\right)^{T}$$

Integrating to find the scaling factor,

$$\begin{aligned} c\int_{\overline{Y}_{MLE}}^{\infty}\left(\frac{1}{\overline{Y}}\right)^{T}d\overline{Y} &= 1 \\ \int_{\overline{Y}_{MLE}}^{\infty}\left(\frac{1}{\overline{Y}}\right)^{T}d\overline{Y} &= \frac{\overline{Y}_{MLE}^{-T+1}}{T-1} \end{aligned}$$

therefore

$$c = (T-1)\left(\overline{Y}_{MLE}\right)^{T-1}$$

Therefore

$$p(\bar{Y}|\{Y_1,\ldots,Y_T\}) = (T-1)\left(\overline{Y}_{MLE}\right)^{T-1}\left(\frac{1}{\overline{Y}}\right)^{T}$$

and

$$\begin{aligned} E[\overline{Y}|\{Y_1,...,Y_T\}] &= c\int_{\overline{Y}_{MLE}}^{\infty}\overline{Y}\left(\frac{1}{\overline{Y}}\right)^{T}d\overline{Y} \\ &= c\int_{\overline{Y}_{MLE}}^{\infty}\left(\frac{1}{\overline{Y}}\right)^{T-1}d\overline{Y} = \frac{c}{T-2}\overline{Y}_{MLE}^{2-T} \\ &= \frac{T-1}{T-2}\overline{Y}_{MLE} \end{aligned}$$

To find the median, note that

$$
\begin{aligned}
c\int_{\overline{Y}_{MLE}}^{med}\left(\frac{1}{\overline{Y}}\right)^{T}d\overline{Y} &= \frac{1}{2}\\
&= (T-1)\,[\overline{Y}_{MLE}]^{T-1}[\frac{(med)^{-T+1}}{-T+1}-\frac{\{\overline{Y}_{MLE}\}^{-T+1}}{-T+1}]\\
\xi_{1/2} &= \left(\frac{1}{2}\right)^{\frac{-1}{T-1}}\overline{Y}_{MLE}
\end{aligned}
$$

(f) Now take the prior to be

$$p\left(\overline{Y}\right)=\frac{1}{\overline{Y}^{2}}$$

Then the posterior will be proportional to

$$\left(\frac{1}{\overline{Y}}\right)^{T+2}$$

to find the scaling factor, note that

$$
\begin{aligned}
c\int_{\overline{Y}_{MLE}}^{\infty}\left(\frac{1}{\overline{Y}}\right)^{T+2}d\overline{Y} &= 1\\
c &= (T+1)\,[\overline{Y}_{MLE}]^{T+1}
\end{aligned}
$$

And the mean is given by

$$c\int_{\overline{Y}_{MLE}}^{\infty}\left(\frac{1}{\overline{Y}}\right)^{T+1}d\overline{Y}=c\frac{\overline{Y}}{-T}^{-T}|_{\overline{Y}_{MLE}}^{\infty}=\frac{T+1}{T}\overline{Y}_{MLE}$$

Calculated as before, the median is given by

$$\xi_{\frac{1}{2}}=\left(\frac{1}{2}\right)^{-\frac{1}{T+1}}\overline{Y}_{MLE}$$

(g) *One approach to answering the admissibility question is to simply note that neither* $\widehat{\log\bar{Y}}$ *nor* $\hat{\bar{Y}}$ *is a function of the sufficient statistic* $Y_{\max}=Y_{MLE}$*. We have cited in class without proof the result that decision rules that are not functions of the sufficient statistic (when one is available) can always be replaced, if loss is convex in the decision rule, by an estimator that has no greater, and possibly lower, loss for every parameter value. The proof is short, so here it is:*

Lemma 0.1 (Jensen's inequality). *If f is a convex function on $\mathbb{R}^n$ and X is a random variable taking values on $\mathbb{R}^n$, then $E[f(X)]>f(E[X])$.*

Lemma 0.2. *If $S(Y)$ is a sufficient statistic for the parameter θ, with the pdf of $Y\mid\theta$ (therefore) taking the form $p(S(y),\theta)g(y)$, then for any function f, with finite expectation, $E[f(Y)\mid S(Y),\theta]=E[f(Y)\mid S(Y)]$, i.e. the expectations conditional on $S(Y)$ do not depend on θ.*

Proposition 0.3. *If $S(Y)$ is a sufficient statistic for θ and expected loss conditional on θ is $E[\mathcal{L}(\theta,\delta(Y))\mid\theta]$ with $\mathcal{L}$ convex, then for every θ, $E\left[\mathcal{L}(\theta,E[\delta(Y)\mid S(Y)])\mid\theta\right]\leq E[\mathcal{L}(\theta,\delta(Y))\mid\theta]$.*

The proposition follows immediately from the two lemmas and the law of iterated expectations.

These two lemmas, though we don't prove them, are facts you should know. Jensen's inequality turns up frequently in economics (though it is more often stated with inequality reversed, for concave functions). The result that conditional expectations given a sufficient statistic don't depend on the parameters is also broadly useful, though more in statistics than economics.

So since the two unbiased estimators both fail to be functions of the sufficient statistic, we can improve both by replacing them by their conditional expectations given $Y_{\max}$*. If we condition on* $Y_1 = Y_{\max}$*, we get, since the* Y_t*'s are independent,*

$$E\Big[\sum_{t=1}^{T} Y_t \mid Y_1 = Y_{\max}\Big] = Y_{\max} + \sum_{t=2}^{T} E[Y_t] = \frac{T+1}{2} Y_{\max}\,,$$

so that $E[\hat{Y} \mid Y_1 = Y_{\max}] = (1 + \frac{1}{T})Y_{\max}$*. The expectation conditional on* $Y_{\max}$*, but not on which observation it matches, is just the equally weighted sum of all the conditional expectations given particular observations equal to* $Y_{\max}$*, and thus, since all are equally likely, is just the same as the expectation conditional on any one of the observations being equal to* $Y_{\max}$*. Note that this is exactly the posterior mean with a* $1/\bar{Y}^2$ *prior.*

The projection of $\widehat{\log Y}$ *on* $Y_{\max}$ *under conditional expectation can be found similarly, and it is* $\log Y_{\max} + \frac{1}{T}$*. This is the posterior mean of* $\log \bar{Y}$ *under an improper* $1/\bar{Y}$ *prior; it is also clearly the limit of posterior means under proper* $\bar{Y}^{-1-\delta}$ *priors as* $\delta \to 0$*.*

Another approach to an answer is to observe that admissible estimators are generally Bayes estimators, or at least limits of Bayes estimators. But Bayes estimators always depend on the data only through the likelihood function, and hence only through the sufficient statistics, so the unbiased estimators are not Bayes for any prior or loss function, hence probably not admissible. Since we haven't given explicit regularity conditions guaranteeing that all admissible estimators are Bayes, you couldn't have been expected to check them.

But, to give a complete answer, having observed how to improve the unbiased estimators by projection on sufficient statistics, you might have proceeded to observe that both of them can be dominated. There is some probability that the mean of the observed Y_t*'s will turn out to be less than* $.5Y_{\max}$*. Indeed the probability of this is about .5. And when this happens, the unbiased estimator* $\hat{Y}$ *is below* $Y_{\max}$*. Since we know that* $\bar{Y} \geq Y_{\max}$*, it always improves the estimator to increase it to* $Y_{\max}$ *whenever it falls below. Similarly it is possible that the mean of the* $\log Y_t$*'s will turn out to be less than* $\log(Y_{\max}) - 1$*, and in such cases it will improve the estimator of* $\log \bar{Y}$ *to replacing it with* $\log(Y_{\max})$*. Since these improvements occur with positive probability, the improved estimators strictly dominate the originals.*