STA 532 HW4 Due Wed 2/12 Yi Mi, ym154

1. (a) $\log \bar{y} = \log \frac{\sum y_i}{n}$

from Jensen's inequality, we have

$$E(f(X)) \geq f(E(X)) \quad \text{for convex } f(\cdot)$$

since $f(x) = \log x$ is concave,

thus,

$$\log \frac{\sum_i y_i}{n} \geq \frac{\sum_i \log y_i}{n} = \frac{\sum_i x_i}{n} = \bar{x}$$

$$\Rightarrow \log \bar{y} \geq \bar{x}$$

$$\bar{y} \geq e^{\bar{x}}$$

(b) let $(z_1 \cdots z_n) = (\frac{1}{y_1} \cdots \frac{1}{y_n})$

so $(x_1 \cdots x_n) = (-\ln z_1 \cdots -\ln z_n)$

from Jensen's inequality, we have

$$\frac{\sum(-\ln z_i)}{n} \geq -\ln \frac{\sum z_i}{n} \quad \text{for convex } f(x) = -\ln x$$

$$\bar{x} \geq -\ln \bar{z}$$

$$e^{\bar{x}} \geq \frac{1}{\bar{z}}$$

using the result from a, we have

$$\bar{y} \geq e^{\bar{x}} \geq \frac{1}{\bar{z}}$$

which means that the magnitude for $f(y) = y$ is greater than that for $f(y) = \ln(y)$, and the later is greater that that for $f(y) = 1/y$

(c) # when $f(y) = \frac{1}{y}$,

$$m(y_1 \cdots y_n) = \frac{n}{\sum \frac{1}{y_i}}$$

and $\frac{\partial}{\partial y_1} m(y_1 \cdots y_n) = \frac{n}{(\sum \frac{1}{y_i})^2} \frac{1}{y_1^2}$

$$\Rightarrow m(y_1+\delta, \cdots y_n) = \frac{n}{\sum \frac{1}{y_i}} + \frac{n}{(\sum \frac{1}{y_i})^2} \frac{1}{y_1^2} \cdot \delta$$

$$= m(y_1 \cdots y_n)\left(1 + \frac{\delta}{y_1^2 \sum \frac{1}{y_i}}\right)$$

As the result, when the $y_1$ is large enoughs the change on $y_1$ will have very less impact, which means it is sensitive with small original values, but not with large original values.

# when $f(y) = \ln y$,

$$m(y_1 \cdots y_n) = (\prod_i y_i)^{\frac{1}{n}}$$

and $\frac{\partial}{\partial y_1} m(y_1 \cdots y_n) = (\prod_i y_i)^{\frac{1}{n}} \frac{1}{n y_1}$

$$\Rightarrow m(y_1+\delta, \cdots y_n) = (\prod_i y_i)^{\frac{1}{n}} + (\prod_i y_i)^{\frac{1}{n}} \frac{1}{n y_1} \cdot \delta$$

$$= m(y_1 \cdots y_n)\left(1 + \frac{\delta}{y_1 n}\right)$$

In this magnitude with outlier, we have $\frac{\delta}{y_1 n}$ term with $n$ in it, meaning that the impact will be shrinked by the sample size, resulting less heavier impact.

# when $f(y) = y$

$$m(y_1 \cdots y_n) = \frac{\sum y_i}{n}$$

and $\frac{\partial}{\partial y_1} m(y_1 \cdots y_n) = \frac{1}{n}$

$$\Rightarrow m(y_1 + \delta, \cdots y_n) = \frac{\sum y_i}{n} + \frac{1}{n} \cdot \delta$$

$$= m(y_1 \cdots y_n)\left(1 + \frac{\delta}{\sum y_i}\right)$$

In this magnitude with outlier, we see that the term $\frac{\delta}{\sum y_i}$ indicates that if the original values are large enough, it will be less sensitive to outliers.

2. (a) $\mathrm{cov}[Y_i, Y_j] = 0$

$$V(\bar{Y}) = V\left(\frac{\Sigma Y_i}{n}\right) = \frac{\Sigma V(Y_i)}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

withe the chebyshev's inequality,

$$0 \le P(|\bar{Y} - \mu| > \varepsilon) \le \frac{\sigma^2}{n} \cdot \frac{1}{\varepsilon^2}$$

$$\lim_{n\to\infty} \frac{\sigma^2}{n} \cdot \frac{1}{\varepsilon^2} = 0$$

thus, $\bar{Y}$ is consistent. for estimating $\mu$.

(b) $\mathrm{cov}[Y_i, Y_j] = \rho$

$$V(\bar{Y}) = V\left(\frac{\Sigma Y_i}{n}\right) = \frac{1}{n^2} Var(\Sigma Y_i)$$

$$= \frac{1}{n^2} E(\Sigma Y_i - E(\Sigma Y_i))^2$$

$$= \frac{1}{n^2} E(\Sigma(Y_i - \mu))^2$$

$$= \frac{1}{n^2} [\Sigma Var(Y_i) + \sum_{i \ne j} \mathrm{cov}(Y_i, Y_j)]$$

$$= \frac{1}{n^2} [n\sigma^2 + n(n-1)\rho]$$

$$= \frac{\sigma^2 + (n-1)\rho}{n}$$

Similarly,

$$0 \le P(|\bar{Y} - \mu| > \varepsilon) \le \left(\frac{\sigma^2}{n} + \left(1 - \frac{1}{n}\right)\rho\right)\frac{1}{\varepsilon^2}$$

$$\lim_{n\to\infty} \left(\frac{\sigma^2}{n} + \left(1 - \frac{1}{n}\right)\rho\right)\frac{1}{\varepsilon^2} = \frac{\rho}{\varepsilon^2}$$

thus, $\bar{Y}$ is not consistent

(c) $\mathrm{cov}[Y_i, Y_j] = \rho$ if $|i-j|=1$ and is zero if $|i-j|>1$

using the derivation above, we have

$$V(\bar{Y}) = \frac{1}{n^2}\left[\sum Var(Y_i) + \sum_{i\neq j} cov(Y_i, Y_j)\right]$$

$$= \frac{1}{n^2}\left[n\sigma^2 + 2(n-1)\rho\right]$$

$$= \frac{n\sigma^2 + 2(n-1)\rho}{n^2}$$

Similarly,

$$0 \le P(|\bar{Y} - \mu| > \varepsilon) \le \frac{n\sigma^2 + 2(n-1)\rho}{n^2} \cdot \frac{1}{\varepsilon^2}$$

$$\lim_{n\to\infty} \frac{n\sigma^2 + 2(n-1)\rho}{n^2} \cdot \frac{1}{\varepsilon^2} = 0$$

thus, $\bar{Y}$ is consistent.

3. (a)

Expectation:

$$E(\hat{\mu}) = E[(1-w)\mu_0 + wY]$$
$$= (1-w)\mu_0 + w\mu$$

Variance:

$$Var(\hat{\mu}) = Var[(1-w)\mu_0 + wY]$$
$$= w^2 Var(Y)$$
$$= w^2\sigma^2$$

bias:

$$B(\hat{\mu}, \mu) = (1-w)\mu_0 + w\mu - \mu$$
$$= (1-w)(\mu_0 - \mu)$$

MSE:

$$MSE(\hat{\mu}, \mu) = Var(\hat{\mu}) + B^2(\hat{\mu}, \mu)$$
$$= w^2\sigma^2 + (1-w)^2(\mu_0 - \mu)^2$$

(b) MSE for $Y$:

$$MSE(Y, \mu) = Var(Y) + B^2(Y, \mu)$$
$$B(Y, \mu) = E(Y) - \mu = 0$$
$$\Rightarrow MSE(Y, \mu) = \sigma^2$$

When $\frac{MSE_{\hat{\mu}}}{MSE_Y} = \frac{w^2\sigma^2 + (1-w)^2(\mu_0 - \mu)^2}{\sigma^2} < 1$,

we have $|\mu - \mu_0| < \sigma$

$$\mu_0 - \sigma < \mu < \mu_0 + \sigma$$

4. With the Chebyshev's inequality, we have

$$\Pr(|\hat{\theta} - \theta| > \varepsilon) \le E[|\hat{\theta} - \theta|^2]/\varepsilon^2$$

and, $MSE(\hat{\theta}, \theta) = E(\hat{\theta} - \theta)^2$

$$\Rightarrow \Pr(|\hat{\theta} - \theta| > \varepsilon) \le \frac{MSE(\hat{\theta}, \theta)}{\varepsilon^2}$$

5. (a) $E(\hat{\mu}_n) = E[(1-w_n)\mu_0 + w_n \bar{Y}_n] = (1-w_n)\mu_0 + w_n\mu$

with Markov's inequality, we have

$$P(|\hat{\mu}_n - \mu| > \varepsilon) \le E[|\hat{\mu}_n - \mu|]/\varepsilon$$

$$= |E\hat{\mu}_n - E\mu|/\varepsilon$$

$$= |(1-w_n)\mu_0 + w_n\mu - \mu|/\varepsilon$$

$$= |(1-w_n)(\mu_0 - \mu)|/\varepsilon$$

when $w_n \to 1$, $|(1-w_n)(\mu_0-\mu)|/\varepsilon \to 0$

and $P(|\hat{\mu}_n - \mu| > \varepsilon) \to 0$ as $n \to \infty$

so the condition is $w_n \to 1$

(b)

i. $P(\mu) \propto N(\mu_0, \tau^2)$

$P(\bar{Y}_n|\mu) \propto N(\mu, \sigma^2/n)$

$$\Rightarrow P(\bar{Y}_n, \mu) = P(\bar{Y}_n|\mu) \times P(\mu)$$

$$= \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left\{-\frac{1}{2}\frac{(\bar{Y}_n-\mu)^2}{\sigma^2/n}\right\} \frac{1}{\sqrt{2\pi\tau^2}} \times \exp\left\{-\frac{1}{2}\frac{(\mu-\mu_0)^2}{\tau^2}\right\}$$

And, $p(\mu \mid \bar{Y}_n) = \dfrac{p(\mu, \bar{Y}_n)}{p(\bar{Y}_n)}$

we know that $p(\bar{Y}_n)$ doesn't depent on $\mu$,

so,

$$p(\mu \mid \bar{Y}_n) \propto p(\mu, \bar{Y}_n)$$

$$\propto \exp\left\{-\frac{1}{2}\left[\frac{\mu^2 - 2\mu\bar{y}_n}{\sigma^2/n} + \frac{\mu^2 - 2\mu \cdot \mu_0}{\tau^2}\right]\right\}$$

let $a = \frac{n}{\sigma^2} + \frac{1}{\tau^2}$

$b = \frac{\bar{y}_n \cdot n}{\sigma^2} + \frac{\mu_0}{\tau^2}$

$$\Rightarrow p(\mu \mid \bar{Y}_n) \propto \exp\left\{-\frac{1}{2}\frac{(\mu - b/a)^2}{1/a}\right\}$$

$$\propto N(b/a, 1/a)$$

with $E(\mu \mid \bar{Y}_n) = b/a = \left(\frac{\bar{y}_n \cdot n}{\sigma^2} + \frac{\mu_0}{\tau^2}\right) / \left(\frac{n}{\sigma^2} + \frac{1}{\tau^2}\right)$

and $Var(\mu \mid \bar{Y}_n) = 1/a = 1 / \left(\frac{n}{\sigma^2} + \frac{1}{\tau^2}\right)$

ii. $P(|\hat{\mu} - \mu| > \varepsilon) \leq E|\hat{\mu} - \mu|^2 / \varepsilon^2$

$$\hat{\mu} = \frac{\frac{\bar{y}_n \cdot n}{\sigma^2} + \frac{\mu_0}{\tau^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}$$

$$E|\hat{\mu} - \mu|^2 / \varepsilon^2 = E\left(\frac{\frac{\bar{y}_n \cdot n}{\sigma^2} + \frac{\mu_0}{\tau^2} - \frac{\mu \cdot n}{\sigma^2} - \frac{\mu}{\tau^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}\right)^2 / \varepsilon^2$$

$=$

$E(\bar{y}_n) = \mu$ $\qquad \mu_0 \qquad \mu_0 \qquad \mu_\infty$

$$E\left|\frac{\frac{n}{\sigma^2}(\bar{y}_n - \mu) + \frac{1}{\tau^2}(\mu_0 - \mu)}{\left(\frac{n}{\sigma^2} + \frac{1}{\tau^2}\right)\varepsilon}\right|$$

$$\frac{\frac{n}{\sigma^2}|\mu - \mu_0|}{\left(\frac{n}{\sigma^2} + \frac{1}{\tau^2}\right)\varepsilon}$$

$$\frac{\frac{1}{\sigma^2}|\mu - \mu_0|}{\left(\frac{1}{\sigma^2} + \frac{1}{n\tau^2}\right)\varepsilon} \to \frac{|\mu - \mu_0|}{\varepsilon}$$