STA532 - hw4 sample solution

Sheng Jiang (sj156@duke.edu)

1\. (a) By Jensen inequality, for some r.v. Z, $\mathbb{E}\log Z \le \log \mathbb{E} Z$.

Let $Z \sim \text{unif}\{y_1, y_2, \cdots, y_n\}$, then $\frac{1}{n}\sum_{i=1}^n \log y_i \le \log \frac{1}{n}\sum_{i=1}^n y_i$, which is $e^{\bar{x}} \le \bar{y}$

(b) Denote $M(\vec{y}, f(\cdot)) = f^{-1}\left(\frac{1}{n}\sum_{i=1}^n f(y_i)\right)$

$f(y) = \log y$ is concave, by Jensen inequality, $\frac{1}{n}\sum_{i=1}^n \log y_i \le \log \frac{1}{n}\sum_{i=1}^n y_i$,

then $M(\vec{y}, \log(\cdot)) \le M(\vec{y}, Id)$, where Id denotes identity map.

Similarly, by Jensen inequality, $\frac{1}{n}\sum_{i=1}^n \log(y_i^{-1}) \le \log \frac{1}{n}\sum_{i=1}^n (y_i^{-1})$,

then $M(\vec{y}, \log(\cdot)) \ge M(\vec{y}, h)$, where $h(x) = \frac{1}{x}$.

To sum up, $M(\vec{y}, h) \le M(\vec{y}, \log(\cdot)) \le M(\vec{y}, Id)$

(c). The three means are differentiable w.r.t. y_1.

$$\frac{\partial}{\partial y_1} M(\vec{y}, Id) = \frac{1}{n}$$

$$\frac{\partial}{\partial y_1} M(\vec{y}, \log(\cdot)) = \frac{\partial}{\partial y_1} e^{\frac{1}{n}\sum_{i=1}^n \log y_i} = \frac{\partial}{\partial y_1} e^{\frac{1}{n}\log y_1} \cdot e^{\frac{1}{n}\sum_{i=2}^n \log y_i} = \frac{1}{n} y_1^{-1} M(\vec{y}, \log(\cdot))$$

$$\frac{\partial}{\partial y_1} M(\vec{y}, h) = \frac{\partial}{\partial y_1} \frac{1}{\frac{1}{n}\sum_{i=1}^n y_i^{-1}} = \frac{1}{(\frac{1}{n}\sum_{i=1}^n y_i^{-1})^2} \frac{1}{n} y_1^{-2} = \frac{1}{n} y_1^{-2} M(\vec{y}, h)^2$$

The contribution of a change in y_1 to arithmic mean is $\frac{1}{n}$ regardless of the value of y_1 or the mean. In contrast, the contribution of a change in y_1 to the other two means *depends* on the value y_1 relative to the mean: if y_1 is smaller than the mean, then the change is greater than $\frac{1}{n}$; vice versa.

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

(a) $V[\bar{Y}] = \frac{1}{n^2}\sum_{i=1}^n V(Y_i) = \frac{1}{n}\sigma^2$

(b) $V[\bar{Y}] = \frac{1}{n^2}\sum_{i=1}^n \sigma^2 + \frac{2}{n^2}\sum_{i<j}\sigma^2\rho = \frac{1}{n}\sigma^2 + \frac{n-1}{n}\rho\sigma^2$

(c) $V[\bar{Y}] = \frac{1}{n^2}\sum_{i=1}^n \sigma^2 + \frac{2}{n^2}\sum_{i<j}\sigma^2\rho \mathbb{1}(|i-j|=1) = \frac{1}{n}\sigma^2 + \frac{2(n-1)}{n^2}\rho\sigma^2$

By Chebyshev inequality, $P(|\bar{Y}-\mu|>\varepsilon) \le \frac{V(\bar{Y})}{\varepsilon^2}$.

For (a), (c), $V(\bar{Y}) \to 0$ as $n\to\infty$. So in (a) & (c), $\bar{Y} \xrightarrow{P} \mu$.

For (b), $V(\bar{Y}) \to \rho\sigma^2 \neq 0$ as $n\to\infty$. Chebyshev inequality does not guarantee convergence in prob for case (b). Note: it's not sufficient to claim inconsistency.

Example: if $\vec{Y} \sim N(\mu 1_n, \rho 1_n^T 1_n + (1-\rho) I_n)$, where $1_n = [\underbrace{1, 1, \cdots 1}_{n \text{ many}}]^T$, I_n is $n\times n$ identity matrix. Then $\bar{Y} \sim N(\mu, V_n)$ with $\lim_{n\to\infty} V_n \neq 0$. In this example, $P(|\bar{Y}_n - \mu| > \varepsilon)$ can be close to 1 for small enough ε and all sufficiently large n.

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

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

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

$MSE(\hat{\mu}) = \text{bias}^2 + V[\hat{\mu}] = (1-w)^2(\mu_0-\mu)^2 + w^2\sigma^2$

(b). $MSE(Y) = \text{bias}^2 + V[Y] = \sigma^2$

$MSE(\hat{\mu}) \le MSE(Y) \iff (1-w)^2(\mu-\mu_0)^2 + w^2\sigma^2 \le \sigma^2$

$\iff (\mu-\mu_0)^2 \le \frac{1+w}{1-w}\sigma^2$

If the introduced bias is small relative to the variance, then the biased estimator is better in terms of MSE.

4. By Chebyshev, $P(|\hat{\theta}-\theta|>\varepsilon) \le \frac{E((\hat{\theta}-\theta)^2)}{\varepsilon^2} = \frac{MSE(\hat{\theta})}{\varepsilon^2}$

5. (a) $\bar{Y}_n \sim N(\mu, \sigma^2/n)$, then by Chebyshev, $\bar{Y}_n \xrightarrow{P} \mu$ as $n\to\infty$.

Suppose the sequence $\{w_n\}$ has a limit w_∞,

then $\hat{\mu}_n \xrightarrow{P} (1-w_\infty)\mu_0 + w_\infty \mu$.

To guarantee $\hat{\mu}_n \xrightarrow{P} \mu$, note limit in probability is unique, "$\lim_{n\to\infty} w_n = 1$" is the minimal condition.

(b) i. $p(\mu, \bar{Y}_n) = p(\bar{Y}_n \mid \mu)\, p(\mu) = \frac{1}{\sqrt{2\pi\sigma^2/n}} e^{-\frac{1}{2}\frac{n}{\sigma^2}(\bar{Y}_n-\mu)^2} \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{1}{2}\tau^{-2}(\mu-\mu_0)^2}$

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

$\propto e^{-\frac{1}{2}(n\sigma^{-2}+\tau^{-2})(\mu-\tilde{\mu})^2}$, where $\tilde{\mu} = \frac{1}{\tau^{-2}+n\sigma^{-2}}(\tau^{-2}\mu_0 + n\sigma^{-2}\bar{Y}_n)$

Clearly, $\mu \mid \bar{Y}_n \sim N(\tilde{\mu}, (\tau^{-2}+n\sigma^{-2})^{-1})$

ii. $\mathbb{E}(\mu \mid \bar{Y}_n) = (\tau^{-2}+n\sigma^{-2})^{-1}(\tau^{-2}\mu_0 + n\sigma^{-2}\bar{Y}_n)$.

$\lim_{n\to\infty} \frac{\tau^{-2}}{\tau^{-2}+n\sigma^{-2}}\mu_0 = 0$, $\bar{Y}_n \xrightarrow{P} \mu$ as $n\to\infty$, $\lim_{n\to\infty}\frac{n\sigma^{-2}}{\tau^{-2}+n\sigma^{-2}} = 1$.

By Slutsky thm, $\mathbb{E}(\mu \mid \bar{Y}_n) \xrightarrow{P} \mu$.