Stat 512
Homework key 6

November 24, 2015

REGULAR PROBLEMS

1. Show that an asymptotically unbiased sequence of estimators need not be consistent. (Hint: Consider independent random variables $X_1, X_2, \ldots$ with $X_i \sim \mathcal{N}(\mu, 2i)$ and let $\bar{X}_n = \sum_{i=1}^n \frac{X_i}{n}$ be an estimator of μ.)

   Ans: Since $E\bar{X}_n = \mu$, it is unbiased, hence asymptotically unbiased. $Var(\bar{X}_n) = 2\sum_{i=1}^n \frac{i}{n^2} = \frac{n+1}{n}$. Hence,

$$\bar{X}_n \sim \mathcal{N}(\mu, \frac{n+1}{n})$$

   Fix $\delta > 0$. Notice that if $X \sim \mathcal{N}(\mu, \sigma^2)$

$$\begin{aligned} P(|X-\theta| \geq \delta) &= P(X-\theta > \delta) + P(X-\theta \leq -\delta) \\ &= 1 - \Phi\left(\frac{\delta+\theta-\mu}{\sigma}\right) + \Phi\left(\frac{-\delta+\theta-\mu}{\sigma}\right) \end{aligned} \tag{1}$$

   In general if $\theta = \mu$,

$$P(|X-\theta| \geq \delta) = 2\left(1 - \Phi\left(\frac{\delta}{\sigma}\right)\right) \tag{2}$$

   Using (2) we get,

$$P(|\bar{X}_n - \mu| > \delta) = 2P(\bar{X}_n - \mu > \delta) = 2\left(1 - \Phi\left(\frac{\delta}{\sqrt{(n+1)/n}}\right)\right) \to 2(1-\Phi(\delta)) \neq 0.$$

   Therefore $\bar{X}_n$is not an consistent estimator of μ.

2. Show that a consistent sequence of estimators for θ need not be asymptotically unbiased.
   (Hint: Consider a sequence of estimators $\{T_n\}_{n=1}^{\infty}$ in which latent mixing random variable $Y_n \sim \mathcal{B}(1, p_n)$ has $(T_n|Y_n = 0) \sim \mathcal{N}(\theta, \frac{\sigma^2}{n})$ and $(T_n|Y_n = 1) \sim \mathcal{N}(n^2, 1)$.)
   Ans:

$$\begin{aligned} &\textbf{if } Y_n = 0,\ T_n = Z_n \sim \mathcal{N}(\theta, \frac{\sigma^2}{n}) \\ &\textbf{if } Y_n = 1,\ T_n = R_n \sim \mathcal{N}(n^2, 1) \end{aligned}$$

   Let $Y_n \sim \mathcal{B}(\frac{1}{n})$. Then,

$$\begin{aligned} &P(|T_n - \theta| \geq \delta) = P(|T_n - \theta| \geq \delta, Y_n = 0) + P(|T_n - \theta| \geq \delta, Y_n = 1) \\ &= P(|T_n - \theta| \geq \delta | Y_n = 0)P(Y_n = 0) + P(|T_n - \theta| \geq \delta | Y_n = 1)P(Y_n = 1) \\ &= P(|Z_n - \theta| \geq \delta)\left(1 - \frac{1}{n}\right) + P(|R_n - \theta| \geq \delta)\frac{1}{n} \\ &= 2\left(1 - \Phi\left(\frac{\delta\sqrt{n}}{\sigma}\right)\right)\left(1 - \frac{1}{n}\right) + \left(1 - \Phi(\delta + \theta - n^2) + \Phi(-\delta + \theta - n^2)\right)\frac{1}{n} \\ &\quad \text{(Using (2) and (1))} \\ &\quad \to 2.0.1 + (1+0+0)0 = 0 \end{aligned}$$

   Hence, T_n is consistent for θ.

$$ET_n = \theta(1 - \frac{1}{n}) + \frac{n^2}{n} \to \infty \tag{3}$$

   Hence, T_n is not asymptotically unbiased.

3. Show that for a sequence of distribution functions $T_n \sim F_{T_n}$ and a limiting distribution function $T \sim F_T$ such that

$$\forall x \in (-\infty, \infty), F_{T_n}(x) \to F_T(x),$$

we can have the "asymptotic variance" $Var(T) \neq \lim_{n\to\infty} Var(T_n)$.
Ans: Consider the set-up of T_n in question 2 with $Z_n = T = N(\theta, 1)$ and $R_n = N(\theta, n^2)$. Then

$$P(T_n = T) = P(Y_n = 1) = 1 - \frac{1}{n} \to 1.$$

Hence, $T_n \xrightarrow{n} T$. Hence, $F_{T_n} \to F_T$. However,

$$\begin{aligned}
&Var(T_n) = Var(E(T_n|Y_n)) + E(Var(T_n|Y_n)) \\
&= E(E(T_n|Y_n)^2) - (E(E(T_n|Y_n))^2 + \frac{1}{n}Var(R_n) + \left(1 - \frac{1}{n}\right)Var(Z_n) \\
&= \frac{1}{n}ER_n^2 + \left(1 - \frac{1}{n}\right)EZ_n^2 - \left(\frac{1}{n}ER_n + \left(1 - \frac{1}{n}\right)EZ_n\right)^2 + \frac{1}{n}Var(R_n) + \left(1 - \frac{1}{n}\right)Var(Z_n) \\
&= \frac{2}{n}Var(R_n) + 2\left(1 - \frac{1}{n}\right)Var(Z_n) + \frac{n-1}{n^2}((ER_n)^2 + (EZ_n)^2) + \frac{2(n-1)}{n^2}ER_nEZ_n
\end{aligned}$$

The first term diverges to ∞ since $Var(R_n) = \infty$ though all other terms are finite since $ER_n, EZ_n, Var(Z_n) < \infty$. Therefore $Var(T_n) \to \infty$ though $Var(T) = 1$.

4. Let $\overrightarrow{X}$ be a random vector and $T(\overrightarrow{X})$ be an estimator of $\theta$with bias function $b(T, \theta)$ show that mean squared error satisfies

$$MSE(T, \theta) = Var(T(\overrightarrow{X})|\theta) + b^2(T, \theta).$$

Ans: Let $ET(\overrightarrow{X}) = \mu$. Hence, $b(T, \theta) = \mu - \theta$. Now,

$$\begin{aligned}
&MSE(T, \theta) \\
&= E(T - \theta)^2 \\
&= E(T - \mu + \mu - \theta)^2 \\
&= E(T - \mu)^2 + E(T - \mu)(\mu - \theta) + (\mu - \theta)^2 \\
&= Var(T(\overrightarrow{X})|\theta) + b^2(T, \theta)
\end{aligned}$$

since the second term is 0. Hence, proved.

5. Let $X_1, X_2, \ldots, X_n$ be a random sample of independent, identically distributed variables with $X_i \sim (\mu, \sigma^2, \omega, \gamma)$, where μ is the first moment, and σ^2, ω and $\gamma < \infty$ are 2nd, 3rd, and 4th central moments, respectively. Define nonparametric estimators of μ and σ^2

$$\begin{aligned}
\bar{X}_n &= \frac{1}{n}\sum_{i=1}^{n} X_i \\
\hat{\sigma}_n^2 &= \frac{1}{n}(X_i - \bar{X}_n)^2 \\
s_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2
\end{aligned}$$

When using a statistic $T(\overrightarrow{X})$ as an estimator of some parameter θ, we define

$$\text{Bias function } b(T(\overrightarrow{X}), \theta) = E[T(\overrightarrow{X})] - \theta$$

$$\text{Mean squared error function } MSE(T(\overrightarrow{X}), \theta) = E[(T(\overrightarrow{X}) - \theta)^2]$$

(a) Find the bias function $b_n(\bar{X}_n, \mu)$ of $\bar{X}_n$ as an estimator of μ. Is $\bar{X}_n$unbiased for μ? Is it asymptotically unbiased for μ?
Ans:

$$E\bar{X}_n = \frac{\sum_{i=1}^{n} EX_i}{n} = \mu$$

Therefore, $\bar{X}_n$unbiased for μ.

(b) Prove or disprove: $\bar{X}_n$is median unbiased for μ in the general case.
Ans: $\bar{X}_n$will be median unbiased for μ if the median of the distribution of $\bar{X}_n$is θ. However, this is not always true, especially for skewed distributions. For example, let $X_1, \ldots, X_n \sim \mathcal{E}(\mu)$ (the mean parameterization). Then $\sum_{i=1}^n X_i \sim \mathcal{E}(n\mu)$. Let us find the median of $Y = \mathcal{E}(\mu)$. Let the median be at M. Then $e^{-\lambda M} = \dfrac{1}{2}$.

$$\begin{aligned}
e^{-M/\mu} &= \frac{1}{2}. \\
\Rightarrow M &= \mu \log 2.
\end{aligned}$$

Therefore, the median of $\bar{X}_n$in this case will be at $\mu \log 2$, not μ. Therefore, $\bar{X}_n$is not median unbiased for μ in the general case.

(c) Find the mean squared error function $MSE(\bar{X}_n, \mu)$ of $\bar{X}_n$ as an estimator of μ.
Ans:

$$\begin{aligned}
&Var(\bar{X}_n) = \frac{\sigma^2}{n}. \\
&b(\bar{X}_n) = 0. \\
&\Rightarrow MSE(\bar{X}_n, \mu) = Var(\bar{X}_n) + b(\bar{X}_n)^2 = \frac{\sigma^2}{n}.
\end{aligned}$$

(d) Find the bias function of $b_n(\hat{\sigma}_n^2, \sigma^2)$ as an estimator of σ^2. Is $\hat{\sigma}_n^2$unbiased for σ^2? Is it asymptotically unbiased for σ^2?
Ans: :

$$\begin{aligned}
&E\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 = E(\sum_{i=1}^{n} X_i^2 - n\bar{X}_n^2) = nEX_i^2 - n(Var(\bar{X}_n) + (E\bar{X}_n)^2) \\
&= n(\mu^2 + \sigma^2) - n(\frac{\sigma^2}{n} + \mu^2) = (n-1)\sigma^2. \\
&\Rightarrow E\hat{\sigma}_n^2 = \frac{n-1}{n}\sigma^2. \\
&\Rightarrow b_n(\hat{\sigma}_n^2, \sigma^2) = E\hat{\sigma}_n^2 - \sigma^2 = -\frac{\sigma^2}{n} \xrightarrow{n} = 0.
\end{aligned}$$

Therefore, $\hat{\sigma}_n^2$ is asymptotically unbiased for σ^2.

(e) Find $Var(\hat{\sigma}_n^2)$ (*Hint: You may find the following results about "quadratic forms" useful:*

- *For random vector $\overrightarrow{Y} = (Y_1, \ldots, Y_n)$ with totally independent components and $Y_i \sim (\mu, \sigma^2, \omega, \gamma < \infty)$, define "quadratic form" $Q = \overrightarrow{Y}^T A \overrightarrow{Y}$ for some n by n matrix A with $diag(A) = \overrightarrow{d}$. Then*

$$Var(Q) = (\gamma - 3\sigma^4)\overrightarrow{d}^T\overrightarrow{d} + 2\sigma^4 tr(AA) + 4\sigma^2\overrightarrow{\mu}^T AA\overrightarrow{\mu} + 4\sigma^2\overrightarrow{\mu}^T A\overrightarrow{d}$$

  $\overrightarrow{\mu} = (\mu_1, \ldots, \mu_n)^T$

- *Consider a quadratic form using $A = I_n - \frac{1}{n}\overrightarrow{1}_n\overrightarrow{1}_n^T$ where I_n is the n by n identity matrix and $\overrightarrow{1}_n$is an n dimensional vector of 1-s.)*

Ans: First we will show that

$$A^2 = AA = A \tag{4}$$

A is called an idempotent matrix.

$$\begin{aligned}
A^2 &= I_n - \frac{2}{n}\overrightarrow{1}_n\overrightarrow{1}_n^T + \frac{\overrightarrow{1}_n^T\overrightarrow{1}_n}{n^2}\overrightarrow{1}_n\overrightarrow{1}_n^T \\
&= I_n - \frac{2}{n}\overrightarrow{1}_n\overrightarrow{1}_n^T + \frac{n}{n^2}\overrightarrow{1}_n\overrightarrow{1}_n^T
\end{aligned}$$

since $\overrightarrow{1}_n^T\overrightarrow{1}_n = n$. Therefore,

$$A^2 = I_n - \frac{1}{n}\overrightarrow{1}_n\overrightarrow{1}_n^T = A$$

Notice $\overrightarrow{a} = Diag(A) = \frac{n-1}{n}\overrightarrow{1_n}$. Also notice that

$$A\overrightarrow{1_n} = 0. \tag{5}$$

Therefore

$$A\overrightarrow{a} = 0 \tag{6}$$

as $A\overrightarrow{1_n} = 0$. Let $\overrightarrow{X} = (X_1, \ldots, X_n)$. Hence,

$$n\hat{\sigma}_n^2 = \overrightarrow{X}^T A\overrightarrow{X}$$

$$\mu = E\overrightarrow{X} = \mu\overrightarrow{1_n}$$

Hence,

$$A\mu = 0 \tag{7}$$

(using (5)).

$$\begin{aligned} Var(\overrightarrow{X}^T A\overrightarrow{X}) &= (\gamma - 3\sigma^4)\overrightarrow{a}^T\overrightarrow{a} + 2\sigma^4 tr(AA) + 4\sigma^2\overrightarrow{\mu}^T AA\overrightarrow{\mu} + 4\sigma^2\overrightarrow{\mu}^T A\overrightarrow{a} \\ =&(\gamma - 3\sigma^4)\overrightarrow{a}^T\overrightarrow{a} + 2\sigma^4 tr(AA) \end{aligned}$$

using (6) and (7). Now $\overrightarrow{a}^T\overrightarrow{a} = \frac{(n-1)^2}{n^2}\overrightarrow{1_n}^T\overrightarrow{1_n} = \frac{(n-1)^2}{n}$. $tr(AA) = tr(A)$ (by (4)). Therefore, $tr(AA) = n - \frac{n}{n} = n - 1$. The variance becomes

$$Var(\overrightarrow{X}^T A\overrightarrow{X}) = \frac{(n-1)^2}{n}(\gamma - 3\sigma^4) + 2(n-1)\sigma^4$$

Hence, $Var(\hat{\sigma}_n^2) = Var(\frac{\overrightarrow{X}^T A\overrightarrow{X}}{n}) = \frac{(n-1)^2}{n^3}(\gamma - 3\sigma^4) + 2\frac{n-1}{n^2}\sigma^4$.

(f) Find the mean squared error function $MSE(\hat{\sigma}_n^2, \sigma^2)$ of $\hat{\sigma}_n^2$as an estimator of σ^2.
Ans:

$$\begin{aligned} MSE(\hat{\sigma}_n^2, \sigma^2) &= Var(\hat{\sigma}_n^2) + b_n^2(\hat{\sigma}_n^2, \sigma^2) \\ &= \frac{(n-1)^2}{n^3}(\gamma - 3\sigma^4) + 2\frac{n-1}{n^2}\sigma^4 + \frac{\sigma^4}{n^2} \\ &= \frac{(n-1)^2}{n^3}(\gamma - 3\sigma^4) + \frac{2n-1}{n^2}\sigma^4 \end{aligned}$$

(g) Find the bias function of $b_n(s_n^2, \sigma^2)$ as an estimator of σ^2. Is s_n^2 unbiased for σ^2? Is it asymptotically unbiased for σ^2?
Ans: :

$$\begin{aligned} s_n^2 &= \frac{n}{n-1}\hat{\sigma}_n^2. \\ \Rightarrow Es_n^2 &= \sigma^2. \end{aligned}$$

Therefore $\hat{\sigma}_n^2$ is unbiased for σ^2.

(h) Find $Var(s_n^2)$.
Ans:

$$\begin{aligned} s_n^2 &= \frac{n}{n-1}\hat{\sigma}_n^2 \\ Var(s_n^2) &= \frac{n^2}{(n-1)^2}Var(\hat{\sigma}_n^2) \\ =&\frac{1}{n}(\gamma - 3\sigma^4) + \frac{2}{(n-1)}\sigma^4. \end{aligned}$$

(i) Find the mean squared error function $MSE(s_n^2, \sigma^2)$ of s_n^2 as an estimator of σ^2.
Ans: Since s_n^2 is unbiased for σ^2, the variance is same as MSE. Therefore $MSE(s_n^2, \sigma^2) = \frac{1}{n}(\gamma - 3\sigma^4) + \frac{2}{(n-1)}\sigma^4$.

(j) Which of the above estimators for σ^2 has smaller MSE?
Ans: x^2 is a convex function of x. Hence, by Jensen's inequality, $EX^2 \geq (EX)^2$. Let $X = (Y_i - EY_i)^2$ $E(Y_i - EY_i)^4 \geq (E(Y - EY_i)^2)^2$. Hence, $\gamma \geq \sigma^4$. Now,

$$\begin{aligned}
&MSE(\hat{\sigma}_n^2) \\
&= \frac{(n-1)^2}{n^3}(\gamma - \sigma^4) - \frac{2(n-1)^2}{n^3}\sigma^4 + \frac{2n-1}{n^2}\sigma^4 \\
&= \frac{(n-1)^2}{n^3}(\gamma - \sigma^4) + \frac{3n-2}{n^3}\sigma^4
\end{aligned}$$

Similarly,

$$\begin{aligned}
&MSE(s_n^2) \\
&= \frac{\gamma - \sigma^4}{n} + \frac{2}{n(n-1)}\sigma^4
\end{aligned}$$

Now,

$$\begin{aligned}
&MSE(s_n^2) - MSE(\hat{\sigma}_n^2) \\
&= \frac{2n-1}{n^3}(\gamma - \sigma^4) - \frac{n^2 - 5n + 2}{n^3(n-1)}\sigma^4 \\
&= \frac{2n-1}{n^3}\gamma - \frac{3n^2 - 8n + 3}{n^3(n-1)}\sigma^4 \\
&= \frac{2n-1}{n^3}(\gamma - \frac{3n^2 - 8n + 3}{2n^2 - 3n + 1}\sigma^4)
\end{aligned}$$

For $n = 1, 2, 3, 4$ we have that

$$\frac{3n^2 - 8n + 3}{2n^2 - 3n + 1} \leq 1.$$

Since $\gamma \geq \sigma^4 \geq p\sigma^4$ (for $0 \leq p \leq 1$), this implies that for $n = 1, 2, 3, 4$, $MSE(s_n^2) \geq MSE(\hat{\sigma}_n^2)$. For $n \geq 5$, the conclusion is dependent on the distribution from which the random sample is drawn. However, for large n, $\frac{3n^2 - 8n + 3}{2n^2 - 3n + 1} \approx \frac{3}{2}$. Therefore for large n, $MSE(s_n^2) > MSE(\hat{\sigma}_n^2)$ if $\gamma > \frac{3}{2}\sigma^4$ and $MSE(s_n^2) \leq MSE(\hat{\sigma}_n^2)$ otherwise.

(k) Show that $s_n = \sqrt{s_n^2}$ is biased for σ.
Ans:

$$\begin{aligned}
Es_n^2 &= Var(s_n) + (Es_n)^2 \\
(Es_n)^2 &= \sigma^2 - Var(s_n) \\
ES_n &= \sqrt{\sigma^2 - Var(s_n)} \leq \sigma.
\end{aligned}$$

In fact unless $Var(s_n) = 0$, i.e. s_n is constant,

$$ES_n = \sqrt{\sigma^2 - Var(s_n)} < \sigma. \tag{8}$$

$Var(s_n) = 0$ only if s_n is a constant random variable, hence X_i is also constant for all i in which case $\sigma = 0$. Except for this trivial case, s_n is biased for σ.

6. Let $X_1, X_2, \ldots$ be a random sample of independent, identically distributed variables with $X_i \sim \mathcal{U}(0, \theta)$. Define estimator $\hat{\theta}_n = max(X_1, X_2, \ldots, X_n)$.

   (a) Show that sequence of estimators $\{\hat{\theta}_n\}_{n=1}^{\infty}$ are asymptotically unbiased and consistent estimators of θ.

   (b) Let $\hat{\theta}_n = a_n\hat{\theta}_n$ be an unbiased estimator of θ. Find a_n, and show that sequence of estimators $\{\hat{\theta}_n\}_{n=1}^{\infty}$ are asymptotically unbiased and consistent estimators of θ.

Ans:

(a) Let $0 < \epsilon < 1$.

$$P(\hat{\theta}_n \leq \theta - \epsilon) = P(X_1, \ldots, X_n \leq \theta - \epsilon) = \left(\frac{\theta - \epsilon}{\theta}\right)^n = \left(1 - \frac{\epsilon}{\theta}\right)^n \to 0$$

$P(\hat{\theta}_n \geq \theta + \epsilon) = 0$. Therefore,

$$P(|\hat{\theta}_n - \theta| \geq \epsilon) \to 0.$$

hence $\hat{\theta}_n$ is consistent for θ. For $0 \leq x \leq \theta$,

$$\begin{aligned} &P(\hat{\theta}_n \leq x) = \left(\frac{x}{\theta}\right)^n \\ &f_{\hat{\theta}_n}(x) = n\frac{x^{n-1}}{\theta^n} \\ &E\hat{\theta}_n = \frac{n}{\theta^n}\int_0^\theta x x^{n-1} dx \\ &= \frac{n}{n+1}\theta. \\ &E\hat{\theta}_n \to \theta \text{ as } n \to \infty. \end{aligned}$$

Hence, $\hat{\theta}_n$ is asymptotically unbiased.

(b)

$$E\frac{n+1}{n}\hat{\theta}_n = \frac{n+1}{n}\frac{n}{n+1}\theta = \theta.$$

Hence, $a_n = \dfrac{n+1}{n}$. $a_n\hat{\theta}_n$ is asymptotically unbiased since it is an unbiased estimator of θ. From part (a) we got that, $\hat{\theta}_n \xrightarrow{P} \theta$. Now $a_n \to 1$. Hence, $a_n \xrightarrow{D} 1$. By Slutskey's lemma, $a_n\hat{\theta}_n \xrightarrow{D} \theta$. Let $\epsilon > 0$.

$$\begin{aligned} &P(|a_n\hat{\theta}_n - \theta| \geq \epsilon) \\ =&P(a_n\hat{\theta}_n \geq \theta + \epsilon) + P(a_n\hat{\theta}_n \leq \theta - \epsilon) \\ =&1 - F_{a_n\hat{\theta}_n}((\theta + \epsilon)-) + F_{a_n\hat{\theta}_n}(\theta - \epsilon) \\ \longrightarrow&1 - 1[\theta + \epsilon > \theta] + 1[\theta - \epsilon \geq \theta](\text{ since } a_n\hat{\theta}_n \xrightarrow{D} \theta) \\ =&1 - 1 + 0 = 0. \end{aligned}$$

Therefore $\{\hat{\theta}_n\}_{n=1}^\infty$ are consistent estimators of θ.

MORE INVOLVED PROBLEMS

7. Let $X_1, X_2, \ldots$ be a random sample of independent, identically distributed variables with $X_i \sim (\mu, \sigma^2 < \infty)$. Chebyshev's inequality provides that for any random variable Y with finite variance and any $\delta > 0$

$$Pr(|Y - E[Y]| \geq \delta) \leq \frac{Var(Y)}{\delta^2}.$$

Use Chebyshev's inequality to prove $\bar{X}_n$ is a consistent estimator of μ.

Ans:

This important result is known as the Weak Law of Large Numbers (WLLN). To prove the WLLN, first note that the sample mean of iid random variables has

$$E\bar{X}_n = \mu, \quad Var\bar{X}_n = \frac{\sigma^2}{n}.$$

Now, let $\epsilon > 0$ be given. By Chebyshev's inequality

$$Pr(|\bar{X}_n - \mu| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2}.$$

Clearly, we have that the bound decreases as $n \to \infty$, converging as

$$b_n \equiv \frac{\sigma^2}{n\epsilon^2} \downarrow 0.$$

Hence, the sequence $Pr(|\bar{X}_n - \mu| \geq \epsilon)$ can be bounded below and above by the sequences $a_n \equiv 0$ and b_n, respectively, with both a_n and b_n converging to 0. This yields the desired result,

$$Pr(|\bar{X}_n - \mu| \geq \epsilon) \to 0,$$

which is equivalent to consistency of the sequence $\bar{X}_n$ for μ. (Note that it suffices for our purposes to state that Chebyshev's inequality has an upper bound that decreases to zero. The remaining details of the argument are typically considered implicit in the setting of convergence in probability.)

8. Let $X_1, X_2, \ldots$ be a random sample of independent, identically distributed variables with $\{T_n(\vec{X})\}_{n=1}^{\infty}$ a sequence of asymptotically unbiased estimators of θ with $Var(T_n) < \infty, \forall n$ and $\lim_{n\to\infty} Var(T_n) = 0$. Show that the sequence of estimators is consistent for θ.

Ans:

Note that if $g(x) \geq 0, \forall x$, then for $r > 0$

$$Eg(X) \geq E[g(X)\mathbf{1}[g(X) \geq r]] \geq rPr(g(X) \geq r).$$

Simple algebra leads us to the inequality

$$Pr(g(X) \geq r) \leq \frac{Eg(X)}{r}.$$

This argument generalizes that of the proof of Chebyshev's inequality. We now let $g(x) = (x - \theta)^2$ and $r = \delta^2 > 0$. We can write

$$Pr([X - \theta]^2 \geq \delta^2) = Pr(|X - \theta| \geq \delta) \leq \frac{E(X - \theta)^2}{\delta^2} = \frac{MSE(X, \theta)}{\delta^2} = \frac{Var(X) + [EX - \theta]^2}{\delta^2}.$$

In the case of this problem, we can consider the sequence $\{T_n(\vec{X})\}_{n=1}^{\infty}$. We have, by a similar argument as in problem 7, that

$$Pr(|T_n(\vec{X}) - \theta| \geq \delta) \leq \frac{Var(T_n(\vec{X})) + [ET_n(\vec{X}) - \theta]^2}{\delta^2} \to 0.$$

Hence, $T_n(\vec{X})$ is consistent for θ. (Note that this result also implies the WLLN since an unbiased estimator is necessarily asymptotically unbiased.)

9. Let $X_1, X_2, \ldots$ be a random sample of independent, identically distributed variables with $X_i \sim \mathcal{B}(1, \theta)$. We are interested in parametric estimators of $Var(X_i) = g(\theta) = \theta(1 - \theta)$.

(a) Show that $T_n(\vec{X}_n) = \bar{X}_n(1-\bar{X}_n)$ is asymptotically unbiased and consistent for $Var(X_i) = g(\theta) = \theta(1-\theta)$.

(b) Find an estimator $T_n^*(\vec{X}_n) = a_n T_n(\vec{X}_n)$ that is unbiased and consistent for $g(\theta)$. How does $MSE(T_n^*, g(\theta))$ compare to $MSE(T_n, g(\theta))$?

(Hint: As derived in class notes, the kurtosis of the Bernoulli distribution is $p(1-p)(1-3p+3p^2) < \infty$.)

Ans:

(a) Since the data are binary, the parametric estimator can be expressedly equivalently as

$$\begin{aligned} T_n(\vec{X}_n) &= \bar{X}_n(1-\bar{X}_n) \\ &= \bar{X}_n - (\bar{X}_n)^2 \\ &= \frac{1}{n}\sum_{i=1}^n X_i - (\bar{X}_n)^2 \\ &= \frac{1}{n}\sum_{i=1}^n X_i^2 - (\bar{X}_n)^2 \\ &= \hat{\sigma}_n^2, \end{aligned}$$

which was defined in problem 5. We know that the bias is

$$-\sigma^2/n \to 0$$

as $n \to \infty$, so T_n is asymptotically unbiased for $g(\theta) = \theta(1-\theta)$. The variance can be written based on the quadratic form Q that was introduced in problem 5. Recall that

$$Var(Q) = \frac{n-1}{n}[(n-1)\gamma - (n-3)\sigma^4].$$

Since $T_n = \hat{\sigma}_n^2 = Q/n$, we again have

$$Var(T_n) = Var(Q/n) = \frac{Var(Q)}{n^2} = \frac{n-1}{n^3}[(n-1)\gamma - (n-3)\sigma^4] \to 0.$$

Applying the result of problem 8 with $\gamma = g(\theta)(1-3g(\theta)) < \infty$ (from the hint/lecture notes) and $\sigma^2 = g(\theta) < \infty$, we find that T_n is consistent for $g(\theta)$.

(b) Using part (a) of this problem along with the results of problem 5, $a_n = n/(n-1)$. Hence,

$$T_n^*(\vec{X}_n) = \frac{n}{n-1}\bar{X}_n(1-\bar{X}_n) = s_n^2.$$

Since $a_n \to_p 1$ and $T_n \to_p g(\theta)$, we obtain that T_n^* is also consistent for $g(\theta)$ by Slutsky's theorem (another argument, based on the result in problem 8, follows along the same lines as part (a) above). By results in problem 5, we know that for large n, the (asymptotic) efficiency depends on the ratio of γ/σ^4,

$$\frac{\gamma}{\sigma^4} = \frac{1-3\theta(1-\theta)}{\theta(1-\theta)} = \frac{1}{\theta(1-\theta)} - 3 > \frac{3}{2} \quad \forall \theta \in (0,1/3) \cup (2/3,1).$$

Hence, for appropriately large values of n, s_n^2 has lower MSE than $\hat{\sigma}_n^2$ if $\theta \in [1/3, 2/3]$.

10. Let $X_1, X_2, \ldots$ be a random sample of independent, identically distributed variables with $X_i \sim \mathcal{P}(\theta)$. We are interested in parametric estimators of $Var(X_i) = g(\theta)$.

(a) Find a parametric estimator $T_n(\vec{X}_n)$ that is unbiased and consistent for $g(\theta)$.

(b) How does the MSE of $T_n(\vec{X}_n)$ compare to the MSE of s_n^2 as found in problem 5?

(Hint: The kurtosis of the Poisson distribution is $3\lambda^2 + \lambda < \infty$.)

Ans:

(a) If $X_i \sim \mathcal{P}(\theta)$, then $E[X_i] = Var(X_i) = \theta$. A natural parametric estimator of $Var(X_i) = \theta$ in this setting is the sample mean $T_n = \bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. We have seen that $\bar{X}_n$ is both unbiased and consistent for $g(\theta) = \theta$, with mean squared error given by

$$MSE(\bar{X}_n, \theta) = Var(\bar{X}_n) + bias(\bar{X}_n, \theta)^2 = \frac{\theta}{n} + 0 \to 0.$$

Applying the result of problem 7 (or 8), we have the desired result.

(b) Unlike in problem 9(b), we need to evaluate the MSE of s_n^2 explicitly to make this comparison. Since s_n^2 was found to be unbiased in problem 5,

$$MSE(s_n^2, \theta) = Var(s_n^2) = Var(Q/(n-1)) = \frac{Var(Q)}{(n-1)^2} = \frac{1}{n(n-1)}[(n-1)\gamma - (n-3)\sigma^4].$$

From the hint, recall that the kurtosis of the Poisson distribution is $\gamma = 3\theta^2 + \theta < \infty$. Furthermore, $\sigma^4 = \theta^2$. Substitution and some algebra yield

$$MSE(s_n^2, \theta) = \frac{\theta}{n} + \frac{2\theta^2}{n-1} \geq \frac{\theta}{n} = MSE(T_n, \theta).$$

In this case, the original parametric estimator T_n is more efficient in terms of MSE.

11. Consider again Chebyshev's inequality.

(a) Show that Chebyshev's inequality can be extremely conservative for $\delta > 1$ by finding a nondegenerate distribution (i.e., with a positive variance) that has 100% of its data within 1 standard deviation of the mean.

(b) Show that Chebyshev's inequality is not always conservative, because for each $\delta \geq 1$ there is a distribution that meets the bound exactly. That is, for arbitrary $\delta \geq 1$, find a distribution F_δ such that exactly $\frac{1}{\delta^2}$ of the data lie at least δ standard deviations away from its mean.

(Hint: Consider a trinomial distribution that takes on values -1 with probability p, 1 with probability p, and 0 with probability 1-2p.)

Ans:

(a) Consider $X \sim \mathcal{B}(1, 1/2)$. Then, $EX = 1/2$, $Var(X) = 1/4$, $SD(X) = 1/2$. Since $X \in \{0, 1\}$, we have constructed a distribution such that 100% of the data is within the interval determined by $1/2 \pm 1/2 \Leftrightarrow [0, 1]$.

(b) Consider a trinomial distribution that takes on values -1 with probability p, 1 with probability p, and 0 with probability $1 - 2p$. It is easy to calculate that this distribution has mean 0 and variance $2p$. For a given $\delta > 1$, we want to choose $p \equiv p(\delta)$ so that

$$Pr(|X - 0| \geq \delta\sqrt{2p}) = \frac{1}{\delta^2}.$$

By choosing $2p = 1/\delta^2$, we obtain $p(\delta) = 1/(2\delta^2)$. The distribution F_δ defined by this choice of trinomial distribution has

$$Pr(|X - 0| \geq \delta\sqrt{2p}) = Pr(|X| \geq 1) = Pr(X \in \{-1, 1\}) = 2p = \frac{1}{\delta^2}.$$