MTH375: Mathematical Statistics - Homework #2

Cason Konzer

Key Concepts: Estimator, bias, unbiased estimator, standard error, MSE, δ-method.

1. Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$be two random samples with the same mean μ and variance σ^2.

(The pdf of X_i and Y_j are not specified.)

(a) Show that $T = \frac{1}{2}\overline{X} + \frac{1}{2}\overline{Y}$ and $U = \frac{1}{3}\overline{X} + \frac{2}{3}\overline{Y}$ are both unbiased estimators of μ.

Solution:

We have $\overline{X} = \dfrac{\sum_{i=1}^{n} X_i}{n}$ and $\overline{Y} = \dfrac{\sum_{j=1}^{n} Y_j}{n}$.

We know for each X_i and Y_j that $E(X_i) = E(Y_j) = \mu$.

It follows that $E(\overline{X}) = E(\overline{Y}) = \dfrac{n \cdot \mu}{n} = \mu$.

Thus $E(T) = \dfrac{1}{2}\mu + \dfrac{1}{2}\mu = \mu$ and $E(U) = \dfrac{1}{3}\mu + \dfrac{2}{3}\mu = \mu$.

It follows that T and U are unbiased estimators of μ as $Bias(T;\mu) = Bias(U;\mu) = \mu - \mu = 0$.

(b) Evaluate $MSE(T;\mu)$ and $MSE(U;\mu)$. According to the MSE criterion, is T or U the better estimator of μ?

Solution:

We know that for unbiased estimators $MSE(S;\theta) = V(S)$.

We also know that for each X_i and Y_j that $V(X_i) = V(Y_j) = \sigma^2$.

It follows that $V(\overline{X}) = V(\overline{Y}) = \dfrac{n \cdot \sigma^2}{n^2} = \dfrac{\sigma^2}{n}$.

We can now compute $V(T)$ and $V(U)$.

$V(T) = V(\dfrac{\overline{X}}{2} + \dfrac{\overline{Y}}{2}) = \dfrac{V(\overline{X})}{4} + \dfrac{V(\overline{Y})}{4} = \dfrac{\sigma^2}{4n} + \dfrac{\sigma^2}{4n} = \dfrac{\sigma^2}{2n}$.

$V(U) = V(\dfrac{\overline{X}}{3} + \dfrac{2\overline{Y}}{3}) = \dfrac{V(\overline{X})}{9} + \dfrac{V(\overline{4Y})}{9} = \dfrac{\sigma^2}{9n} + \dfrac{4\sigma^2}{9n} = \dfrac{5\sigma^2}{9n}$.

Thus $MSE(T;\mu) = \dfrac{9\sigma^2}{18n}$ and $MSE(U;\mu) = \dfrac{10\sigma^2}{18n}$.

We can now see that T is a better estimator of μ than U for all sample sizes as the MSE is closer to 0 for all sample sizes n; both are consistent estimators.

2. Let $X_1, \ldots, X_n$ be a random sample uniform on $[0, \theta]$.
(Hint: We know f_{X_i}, $E(X_i)$ and $V(X_i)$; use them all.)

(a) Show that $T = 2\overline{X}$ is an unbiased estimator of θ, and evaluate $MSE(T;\theta)$.

Solution:

We know that $f_{X_i} = \dfrac{1}{\theta - 0} = \dfrac{1}{\theta}$, $E(X_i) = \dfrac{0+\theta}{2} = \dfrac{\theta}{2}$, $V(X_i) = \dfrac{(\theta-0)^2}{12} = \dfrac{\theta^2}{12}$.

We also know that $\overline{X} = \dfrac{\sum_{i=1}^n X_i}{n}$.

We can evaluate $E(\overline{X}) = \dfrac{\theta n}{2n} = \dfrac{\theta}{2}$.

It follows nicely that $E(T) = E(2\overline{X}) = 2E(\overline{X}) = \dfrac{2\theta n}{2n} = \theta$.

We can now show that T is an unbiased estimator of θ as $Bias(T;\theta) = \theta - \theta = 0$.

As T is an unbiased of θ, we know that $MSE(T;\theta) = V(T)$.

We can evaluate $V(\overline{X}) = \dfrac{nV(X_i)}{n^2} = \dfrac{\theta^2}{12n}$.

It follows that $V(T) = V(2\overline{X}) = 4V(\overline{X}) = \dfrac{4\theta^2}{12n} = \dfrac{\theta^2}{3n}$.

We can now see that $MSE(T;\theta) = \dfrac{\theta^2}{3n}$ and that T is a consistent estimator of θ.

(b) Let $M = max\{X_1, ..., X_n\}$. Find the *pdf* of M.
(Hint: Use the fact that $F_m = P(M \le m) = P(X_1 \le m \,\&\cdots\&\, X_n \le m)$.)

Solution:

We know that $P(M \le m) = P(X_1 \le m \,\&\cdots\&\, X_n \le m) = (F_{X_i})^n$.

We also know that $F_{X_i} = \displaystyle\int_0^x \frac{1}{\theta}\,dx = \frac{x}{\theta}\Big|_0^x = \frac{x}{\theta}$.

It then follows that $F_m = \left(\dfrac{x}{\theta}\right)^n = \dfrac{x^n}{\theta^n}$.

We can now see that $f_m = \dfrac{d}{dx}[F_m] = \dfrac{d}{dx}\left[\dfrac{x^n}{\theta^n}\right] = \dfrac{nx^{n-1}}{\theta^n}$.

(c) Compute $E(M)$ and $V(M)$.

Solution:

From probability we know that $E(M) = \int_0^\theta x f_m(x)\,dx = \int_0^\theta \frac{nx^n}{\theta^n}\,dx = \frac{n}{\theta^n}\int_0^\theta x^n\,dx.$

Thus $E(M) = \frac{n}{\theta^n}\left[\frac{x^{n+1}}{n+1}\Big|_0^\theta\right] = \frac{n}{\theta^n}\left[\frac{\theta^{n+1}}{n+1} - \frac{0^{n+1}}{n+1}\right] = \frac{n\theta^{n+1}}{(n+1)\theta^n} = \frac{n\theta}{n+1}.$

We will now solve for $V(M) = E(M^2) - E(M)^2$.

It is straight forward that $E(M)^2 = \frac{n^2\theta^2}{(n+1)^2}$.

By RUS, $E(M^2) = \int_0^\theta x^2 f_m(x)\,dx = \int_0^\theta \frac{nx^{n+1}}{\theta^n} = \frac{n}{\theta^n}\int_0^\theta x^{n+1}\,dx.$

Thus $E(M) = \frac{n}{\theta^n}\left[\frac{x^{n+2}}{n+2}\Big|_0^\theta\right] = \frac{n}{\theta^n}\left[\frac{\theta^{n+2}}{n+2} - \frac{0^{n+2}}{n+2}\right] = \frac{n\theta^{n+2}}{(n+2)\theta^n} = \frac{n\theta^2}{n+2}.$

We now have $V(M) = \frac{n\theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2} = \frac{n\theta^2(n+1)^2 - n^2\theta^2(n+2)}{(n+2)(n+1)^2}$

$= \frac{n\theta^2(n^2+2n+1) - (n^3\theta^2 + 2n^2\theta^2)}{(n+2)(n+1)^2} = \frac{n^3\theta^2 + 2n^2\theta^2 + n\theta^2 - n^3\theta^2 - 2n^2\theta^2}{(n+2)(n+1)^2}.$

$= \frac{n\theta^2}{(n+2)(n+1)^2}.$

(d) Using your answers to (c), find an unbiased estimator M^* of θ based on M, and evaluate $MSE(M^*;\theta)$. According to the MSE criterion, which is a better estimator of θ, T or M^*?

Solution:

We can obtain an unbiased estimator M^* of θ by solving $E(M) = \frac{n\theta}{n+1}$ for θ.

$$\theta = \frac{(n+1)E(M)}{n} = E\Big(\frac{(n+1)M}{n}\Big) = E\left(\frac{(n+1)max\{X_1,...,X_n\}}{n}\right).$$

We now have an unbiased estimator $M^* = \frac{(n+1)max\{X_1,...,X_n\}}{n}$ for θ.

Now we will solve for $MSE(M^*) = V(M^*) = V\Big(\frac{(n+1)M}{n}\Big) = \frac{(n+1)^2}{n^2}V(M).$

$= \frac{(n+1)^2}{n^2}\cdot\frac{n\theta^2}{(n+2)(n+1)^2} = \frac{\theta^2}{n(n+2)}.$

As $\lim_{n\to\infty}\frac{\theta^2}{n(n+2)} = 0$, We can say that M^* is a consistent estimator of θ.

Last, as $\dfrac{\theta^2}{n(n+2)} = \dfrac{\theta^2}{n^2+2n} < \dfrac{\theta^2}{3n}$ for all $n > 1$,

we find that M^* is a better estimator of θ than T.

---

3. Let $X_1, \ldots, X_n$ be a sample of *iid* $Bin(1, \theta)$ random variables, and let $T = \overline{X}(1 - \overline{X})$ be an estimator of $V(X_i) = \theta(1-\theta)$.

(Hint: We know f_{X_i}, $E(X_i)$ and $V(X_i)$; use them all.)

$$f_{X_i} = \binom{1}{x}\theta^x(1-\theta)^{1-x} \quad ; \quad E(X_i) = \theta \quad ; \quad V(X_i) = \theta(1-\theta).$$

(a) Determine $E(T)$.

Solution:

We will first let $\displaystyle\sum_{i=1}^{n} X_i = \Sigma$ as shorthand.

We can see that $\Sigma \sim Bin(n, \theta)$.

It follows that $E(\Sigma) = n\theta$ and $V(\Sigma) = n\theta(1-\theta)$.

We can now solve for $E(\overline{X}) = \dfrac{E(\Sigma)}{n} = \theta$ and $V(\overline{X}) = \dfrac{V(\Sigma)}{n^2} = \dfrac{\theta(1-\theta)}{n}$.

We know $V(\overline{X}) = E(\overline{X}^2) - E(\overline{X})^2$ and $E(\overline{X}^2) = V(\overline{X}) + E(\overline{X})^2$.

It is straightforward that $E(\overline{X})^2 = \theta^2$.

We will now solve for $E(\overline{X}^2) = \dfrac{V(\Sigma)}{n^2} = \dfrac{\theta(1-\theta)}{n} + \theta^2$.

We can now see that $E(T) = E(\overline{X}(1-\overline{X})) = E(\overline{X} - \overline{X}^2) = E(\overline{X}) - E(\overline{X}^2)$.

By substitution, $E(T) = \theta - \dfrac{\theta(1-\theta)}{n} - \theta^2 = \dfrac{n\theta - \theta + \theta^2 - n\theta^2}{n} = \dfrac{\theta(n-1) - \theta^2(n-1)}{n}$.

Furter simplification arrives at $E(T) = \dfrac{\theta(1-\theta)(n-1)}{n}$.

(b) Determine $Bias(T; \theta(1-\theta))$.

Solution:

$Bias(T; \theta(1-\theta)) = E(T) - \theta(1-\theta) = \dfrac{\theta(1-\theta)(n-1)}{n} - \theta(1-\theta) = \dfrac{\theta(1-\theta)(n-1) - n\theta(1-\theta)}{n}$.

Simplification provides our solution, $Bias(T; \theta(1-\theta)) = \dfrac{\theta(1-\theta)(n-1-n)}{n} = -\dfrac{\theta(1-\theta)}{n}$.

(c) Determine the asymptotic bias of T for estimating $\theta(1-\theta)$.

Solution:

We can see that T is an asymptotically unbiased estimator for $\theta(1-\theta)$ as $\lim_{n\to\infty} -\dfrac{\theta(1-\theta)}{n} = 0$.

(d) Determine an unbiased estimator of $\theta(1-\theta)$ based on T.

Solution:

As $E(T) = \dfrac{\theta(1-\theta)(n-1)}{n}$ We can solve for $\theta(1-\theta)$ to find an unbiased estimator.

$$\theta(1-\theta) = \frac{nE(T)}{n-1} = E\Big(\frac{nT}{n-1}\Big) = E\Big(\frac{n\overline{X}(1-\overline{X})}{n-1}\Big).$$

We can now see that $\dfrac{n\overline{X}(1-\overline{X})}{n-1}$ is an unbiased estimator of $\theta(1-\theta)$.

---

4. Let $X_1, \ldots, X_n$ be a sample of *iid* $N(0, \sigma^2)$ random variables, and let $T = \dfrac{1}{n}\sum_{i=1}^{n} X_i^2$. (Hint: the mgf of X_i is $M(t) = e^{\sigma^2 t^2/2}$.)

(a) Is T an unbiased estimator of σ^2?

Solution:

We know that $V(X_i) = E(X_i^2) - E(X_i)^2$ and that $E(X_i) = 0$.

We can then see that $E(X_i^2) = V(X_i) + E(X_i)^2 = \sigma^2 + 0^2 = \sigma^2$.

It follows nicely that $E(T) = \dfrac{nE(X_i^2)}{n} = \sigma^2$.

We can now see that T is an unbiased estimator of σ^2 as $Bias(T;\sigma^2) = \sigma^2 - \sigma^2 = 0$.

(b) Find $MSE(T;\sigma^2)$.

Solution:

We know $MSE(T;\sigma^2) = V(T)$. as T is unbiased.

Notice $V(T) = V\Big(\dfrac{1}{n}\sum_{i=1}^{n} X_i^2\Big) = \dfrac{1}{n^2}\sum_{i=1}^{n} V(X_i^2) = \dfrac{1}{n^2}\sum_{i=1}^{n} E(X_i^4) - E(X_i^2)^2 = \dfrac{E(X_i^4) - E(X_i^2)^2}{n}$

Differentiating the mgf, $M(t) = e^{\sigma^2 t^2/2}$; $M'(t) = \sigma^2 t e^{\sigma^2 t^2/2}$; $M^2(t) = \sigma^2 e^{\sigma^2 t^2/2} + \sigma^4 t^2 e^{\sigma^2 t^2/2}$.

$$M(t)^3 = \sigma^4 t e^{\sigma^2 t^2/2} + 2\sigma^4 t e^{\sigma^2 t^2/2} + \sigma^6 t^3 e^{\sigma^2 t^2/2} = 3\sigma^4 t e^{\sigma^2 t^2/2} + \sigma^6 t^3 e^{\sigma^2 t^2/2}$$

$$M(t)^4 = 3\sigma^4 e^{\sigma^2 t^2/2} + 3\sigma^6 t^2 e^{\sigma^2 t^2/2} + 3\sigma^6 t^2 e^{\sigma^2 t^2/2} + \sigma^8 t^4 e^{\sigma^2 t^2/2}$$

$$M(t)^4 = 3\sigma^4 e^{\sigma^2 t^2/2} + 6\sigma^6 t^2 e^{\sigma^2 t^2/2} + \sigma^8 t^4 e^{\sigma^2 t^2/2}$$

$$M(0)^4 = 3\sigma^4 e^{\sigma^2 (0)^2/2} + 6\sigma^6 (0)^2 e^{\sigma^2 (0)^2/2} + \sigma^8 (0)^4 e^{\sigma^2 (0)^2/2} = 3\sigma^4 = E(X_i^4).$$

From above, $E(X_i^2) = \sigma^2$, Thus $E(X_i^2)^2 = (\sigma^2)^2 = \sigma^4$.

We can now solve $MSE(T;\sigma^2) = V(T) = \dfrac{E(X_i^4) - E(X_i^2)^2}{n} = \dfrac{3\sigma^4 - \sigma^4}{n} = \dfrac{2\sigma^4}{n}$.

In conclusion we say that T is a consistent estimator of σ^2 as $\lim\limits_{n\to\infty} \dfrac{2\sigma^4}{n} = 0$.

---

5. Let $X_1, \ldots, X_n$ be a sample of *iid* $Exp(\beta)$ random variables. Use the δ-Method to determine the approximate standard error of $\hat{\beta}^2 = \overline{X}^2$.

Solution:

Recall the δ-Method, given a function g and $SE(S)$, $SE(g(S)) \approx |g'(E(S))| \cdot SE(S)$.

Here we will apply the δ-Method for $S = \overline{X}$; $g(S) = S^2$; $g'(S) = 2S$.

For shorthand we will us $\Sigma = \sum_{i=1}^{n} X_i$ as $\Sigma \sim Gamma(n, \beta)$.

It follows that $E(\overline{X}) = E\left(\frac{\Sigma}{n}\right) = \frac{n\beta}{n} = \beta$.

Similarly $V(\overline{X}) = V\left(\frac{\Sigma}{n}\right) = \frac{n\beta^2}{n^2} = \frac{\beta^2}{n}$; $SE(\overline{X}) = \sqrt{V(\overline{X})} = \sqrt{\frac{\beta^2}{n}} = \frac{\beta}{\sqrt{n}}$.

Now $SE(\hat{\beta}^2) = SE(g(\overline{X})) \approx |g'(E(\overline{X}))| \cdot SE(\overline{X}) = |g'(\beta)| \cdot \frac{\beta^2}{\sqrt{n}} = |2\beta| \cdot \frac{\beta}{\sqrt{n}}$.

We now have an approximate of $SE(\hat{\beta}^2)$, given as $SE(\hat{\beta}^2) \approx \frac{2\beta}{\sqrt{n}}$.