Biost/Stat 533
Spring 2008

Homework Assignment #2

Problem 1

Prove if $\boldsymbol{a}$ is a vector of constants with the same dimension of the random vector $\boldsymbol{X}$, then

$$E[(\boldsymbol{X}-\boldsymbol{a})(\boldsymbol{X}-\boldsymbol{a})^T] = Var[\boldsymbol{X}] + (E[\boldsymbol{X}]-\boldsymbol{a})(E[\boldsymbol{X}]-\boldsymbol{a})^T.$$

If $Var[\boldsymbol{X}] = (\sigma_{ij})$, show

$$E[\|\boldsymbol{X}-\boldsymbol{a}\|^2] = \sum_i \sigma_{ij} + \|E[\boldsymbol{X}]-\boldsymbol{a}]\|^2.$$

Solution: Let $E[\boldsymbol{X}] = \mu$. We have

$$\begin{aligned}
E[(\boldsymbol{X}-\boldsymbol{a})(\boldsymbol{X}-\boldsymbol{a})^T] &= E[(\boldsymbol{X}-\boldsymbol{\mu}+\boldsymbol{\mu}-\boldsymbol{a})(\boldsymbol{X}-\boldsymbol{\mu}+\boldsymbol{\mu}-\boldsymbol{a})^T] \\
&= E[(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})^T + (\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{\mu}-\boldsymbol{a})^T + (\boldsymbol{\mu}-\boldsymbol{a})(\boldsymbol{X}-\boldsymbol{\mu})^T + (\boldsymbol{\mu}-\boldsymbol{a})(\boldsymbol{\mu}-\boldsymbol{a})^T] \\
&= E[(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})^T] + E[\boldsymbol{X}-\boldsymbol{\mu}](\boldsymbol{\mu}-\boldsymbol{a})^T + (\boldsymbol{\mu}-\boldsymbol{a})E[(\boldsymbol{X}-\boldsymbol{\mu})^T] + (\boldsymbol{\mu}-\boldsymbol{a})(\boldsymbol{\mu}-\boldsymbol{a})^T \\
&= Var[\boldsymbol{X}] + 0 + 0 + (\boldsymbol{\mu}-\boldsymbol{a})(\boldsymbol{\mu}-\boldsymbol{a})^T \\
&= Var[\boldsymbol{X}] + (E[\boldsymbol{X}]-\boldsymbol{a})(E[\boldsymbol{X}]-\boldsymbol{a})^T.
\end{aligned}$$

Note that the trace of a scalar is a scalar itself. Using this and applying the previous result, we have

$$\begin{aligned}
E[\|\boldsymbol{X}-\boldsymbol{a}\|^2] &= E[\text{trace}(\|\boldsymbol{X}-\boldsymbol{a}\|^2)] \\
&= E[\text{trace}((\boldsymbol{X}-\boldsymbol{a})^T(\boldsymbol{X}-\boldsymbol{a}))] \\
&= E[\text{trace}((\boldsymbol{X}-\boldsymbol{a})(\boldsymbol{X}-\boldsymbol{a})^T)] \\
&= \text{trace}(E[(\boldsymbol{X}-\boldsymbol{a})(\boldsymbol{X}-\boldsymbol{a})^T]) \\
&= \text{trace}(Var[\boldsymbol{X}] + (E[\boldsymbol{X}]-\boldsymbol{a})(E[\boldsymbol{X}]-\boldsymbol{a})^T) \\
&= \text{trace}(Var[\boldsymbol{X}]) + \text{trace}((E[\boldsymbol{X}]-\boldsymbol{a})(E[\boldsymbol{X}]-\boldsymbol{a})^T)) \\
&= \sum_i \sigma_{ij} + \text{trace}((E[\boldsymbol{X}]-\boldsymbol{a})^T E[\boldsymbol{X}]-\boldsymbol{a})) \\
&= \sum_i \sigma_{ij} + \text{trace}(\|E[\boldsymbol{X}]-\boldsymbol{a}]\|^2) \\
&= \sum_i \sigma_{ij} + \|E[\boldsymbol{X}]-\boldsymbol{a}]\|^2.
\end{aligned}$$

Problem 2

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)^T$ be a vector of random variables and let $Y_1 = X_1$, $Y_i = X_i - X_{i-1} (i = 2, 3, \ldots, n)$. If Y_i are mutually independent random variables each with unit variance, find $Var(\boldsymbol{X})$.

Solution: We have that $X_i = \sum_{n=1}^{i} Y_n$, or in matrix notation, $\boldsymbol{X} = A\boldsymbol{Y}$, where

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix}$$

and

$$Var[\boldsymbol{X}] = Var[A\boldsymbol{Y}] = AVar[\boldsymbol{Y}]A^T = AIA^T = AA^T$$

Therefore

$$Var[\boldsymbol{X}] = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & n \end{bmatrix}.$$

Problem 3

If $X_1, X_2, \ldots, X_n$ are random variables satisfying $X_{i+1} = \rho X_i$ where ρ is a constant, and $Var(X_1) = \sigma^2$, find $Var(\boldsymbol{X})$.

Solution: Let

$$\boldsymbol{a} = \begin{bmatrix} 1 \\ \rho \\ \rho^2 \\ \vdots \\ \rho^{n-1} \end{bmatrix}.$$

Then $\boldsymbol{X} = \boldsymbol{a}X_1$ and we have

$$\begin{aligned} Var(\boldsymbol{X}) &= Var(\boldsymbol{a}X_1) \\ &= \boldsymbol{a}Var(X_1)\boldsymbol{a}^T \\ &= \sigma^2 \boldsymbol{a}\boldsymbol{a}^T \\ &= \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & \rho^2 & \rho^3 & \cdots & \rho^n \\ \rho^2 & \rho^3 & \rho^3 & \cdots & \rho^{n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^n & \rho^{n+1} & \cdots & \rho^{2n-2} \end{bmatrix}. \end{aligned}$$

Problem 4

If $X_1, X_2, \ldots, X_n$ are independent random variables with common mean μ and variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$, find $Var[\bar{X}]$.

Solution:

$$\begin{aligned} Var[\bar{X}] &= Var[\frac{1}{n}\sum_{i=1}^n X_i] \\ &= \frac{1}{n^2} Var[\sum_{i=1}^n X_i] \\ &= \frac{\sum_{i=1}^n \sigma_i^2}{n^2}. \end{aligned}$$

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)^T$ and $\boldsymbol{1_n} \in \mathbb{R}^n$ be the vector whose elements are all 1. Define $\bar{J_n} = \boldsymbol{1_n}\boldsymbol{1_n}^T/n$. Recall from the lecture note that

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \boldsymbol{X}^T A \boldsymbol{X}$$

where

$$A = I_n - \bar{J_n}$$

and that

$$A\boldsymbol{1_n} = 0.$$

Therefore

$$\begin{aligned} E[\sum_{i=1}^n (X_i - \bar{X})^2] &= E[\boldsymbol{X}^T A \boldsymbol{X}] \\ &= \text{trace}(ACov(\boldsymbol{X})) + (\mu\boldsymbol{1_n})^T A(\mu\boldsymbol{1_n}) \\ &= \text{trace}(A\text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)) + 0 \\ &= \frac{n-1}{n}\sum_{i=1}^n \sigma_i^2 \end{aligned}$$

and

$$\begin{aligned} E[\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n(n-1)}] &= \frac{1}{n(n-1)} \frac{n-1}{n} \sum_{i=1}^n \sigma_i^2 \\ &= \frac{\sum_{i=1}^n \sigma_i^2}{n^2} \\ &= Var[\bar{X}] \end{aligned}$$

Problem 5

Let $X_1, X_2, \ldots, X_n$ be independently and identically distributed as $N(\theta, \sigma^2)$. Define

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

and

$$Q = \frac{1}{2(n-1)}\sum_{i=1}^{n}(X_{i+1} - X_i)^2.$$

(a)

Prove that $var[S^2] = 2\sigma^4/(n-1)$.

(b)

Show that Q is an unbiased estimator of σ^2.

(c)

Find the variance of Q and show that as $n \to \infty$ the efficiency of Q relative to S^2 is $\frac{2}{3}$.

Solution:

1. Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)^T$ and $\boldsymbol{1_n} \in \mathbb{R}^n$ be the vector whose elements are all 1. Define $\bar{J_n} = \boldsymbol{1_n}\boldsymbol{1_n^T}/n$ and $\boldsymbol{\theta} = \theta\boldsymbol{1_n}$.

$$\sum_{i=1}^{n}(X_i - \bar{X})^2 = \boldsymbol{X}^T A \boldsymbol{X}$$

where

$$A = I_n - \bar{J_n}.$$

Because A is a symmetric matrix and $X_1, X_2, \ldots, X_n$ is an IID sample, we can apply Theorem 1.6 that

$$var[\boldsymbol{X}^T A \boldsymbol{X}] = (\mu_4 - 3\mu_2^2)\boldsymbol{a}^T\boldsymbol{a} + 2\mu_2^2\text{trace}(A^2) + 4\mu_2\boldsymbol{\theta}^T A^2\boldsymbol{\theta} + 4\mu_3\boldsymbol{\theta}^T A\boldsymbol{a}$$

where μ_i represents the ith central moment of the X_i's and $\boldsymbol{a}$ is a vector of the diagonal elements of A. For the normal distribution

$$\mu_3 = 0$$

and

$$\mu_4 = 3\mu_2^2 = 3\sigma^4$$

Therefore

$$\begin{aligned}
var[\sum_{i=1}^{n}(X_i - \bar{X})^2] &= (\mu_4 - 3\mu_2^2)\boldsymbol{a}^T\boldsymbol{a} + 2\mu_2^2\text{trace}(A^2) + 4\mu_2\boldsymbol{\theta}^T A^2\boldsymbol{\theta} + 4\mu_3\boldsymbol{\theta}^T A\boldsymbol{a} \\
&= (3\mu_2^2 - 3\mu_2^2)\boldsymbol{a}^T\boldsymbol{a} + 2\mu_2^2\text{trace}((I_n - \bar{J}_n)^2) + 4\mu_2\boldsymbol{\theta}^T(I_n - \bar{J}_n)^2\boldsymbol{\theta} + 4\times 0\times\boldsymbol{\theta}^T A\boldsymbol{a} \\
&= 2\mu_2^2\text{trace}(I_n - \bar{J}_n) \\
&= 2\sigma^4(n-1)
\end{aligned}$$

Note that $\boldsymbol{\theta}^T(I_n - \bar{J}_n)^2\boldsymbol{\theta} = \boldsymbol{\theta}^T(I_n - \bar{J}_n)\boldsymbol{\theta} = 0$. Therefore

$$\begin{aligned}
var[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2] &= \frac{1}{(n-1)^2}var[\sum_{i=1}^{n}(X_i - \bar{X})^2] \\
&= \frac{1}{(n-1)^2}2\sigma^4(n-1) \\
&= \frac{2\sigma^4}{n-1}
\end{aligned}$$

2.

$$\begin{aligned}
E[Q] &= E[\frac{1}{2(n-1)}\sum_{i=1}^{n-1}(X_{i+1} - X_i)^2] \\
&= \frac{1}{2(n-1)}E[\sum_{i=1}^{n-1}(X_{i+1}^2 - 2X_{i+1}X_i + X_i^2)] \\
&= \frac{1}{2(n-1)}\sum_{i=1}^{n-1}(\mu^2 + \sigma^2 - 2\mu^2 + \mu^2 + \sigma^2) \\
&= \frac{1}{2(n-1)}\sum_{i=1}^{n-1}(2\sigma^2) \\
&= \sigma^2
\end{aligned}$$

3. Let $\boldsymbol{Y} = (Y_1, Y-2, \ldots, Y_n)^T$ where $Y_i = X_{i+1} - X_i$ for $i = 1, \cdots, n-1$ and $Y_n = 0$. Then $\boldsymbol{Y} = A\boldsymbol{X}$ where

$$A = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ 0 & 0 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Therefore

$$\sum_{i=1}^{n-1}(X_{i+1} - X_i)^2 = \boldsymbol{X}' A^* \boldsymbol{X}$$

where

$$A^* = A'A = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}$$

and

$$A^{*2} = \begin{bmatrix} 2 & -3 & 1 & 0 & \cdots & 0 & 0 \\ -3 & 6 & -4 & 1 & \cdots & 0 & 0 \\ 1 & -4 & 6 & -4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 6 & -3 \\ 0 & 0 & 0 & 0 & \cdots & -3 & 2 \end{bmatrix}$$

Since A^* is a symmetric matrix, we can use the same variance formula as in part (a)

$$var[\sum_{i=1}^{n-1}(X_{i+1} - X_i)^2] \quad = \quad (\mu_4 - 3\mu_2^2)\boldsymbol{a}^{*T}\boldsymbol{a}^* + 2\mu_2^2 \text{trace}(A^{*2}) + 4\mu_2 \boldsymbol{\theta}^T A^{*2} \boldsymbol{\theta} + 4\mu_3 \boldsymbol{\theta}^T A^* \boldsymbol{a}^*$$

where $\boldsymbol{a}^*$ is a vector of the diagonal elements of A^*,

$$(\mu_4 - 3\mu_2^2)\boldsymbol{a}^{*'}\boldsymbol{a}^* = 0$$

and

$$4\mu_3 \boldsymbol{\theta}^T A^* \boldsymbol{a}^*$$

Also

$$\boldsymbol{\theta}^T A^* = \theta \mathbf{1}_{\boldsymbol{n}}^T A^* = 0$$

because the columns of A^* sum to 0. Therefore

$$\begin{aligned} var[\sum_{i=1}^{n-1}(X_{i+1} - X_i)^2] &= 2\mu_2^2 \text{trace}(A^{*2}) \\ &= 2\sigma^4(6(n-2)+4) \\ &= 2\sigma^4(6n-8) \end{aligned}$$

and

$$var[Q] = \frac{\sigma^4(6n-8)}{2(n-1)^2}$$

We see that

$$\begin{aligned} \frac{var[S^2]}{var[Q]} &= \frac{\frac{2\sigma^4}{n-1}}{\frac{\sigma^4(6n-8)}{2(n-1)^2}} \\ &= \frac{4n-4}{6n-8} \\ &\rightarrow \frac{2}{3} \end{aligned}$$

asymptotically, Q is $\frac{2}{3}$ as efficient as S^2.

Problem 6

Let $\boldsymbol{X} = (X_1, X_2, X_3)^T$ with

$$Var[\boldsymbol{X}] = \begin{bmatrix} 5 & 2 & 3 \\ 2 & 3 & 0 \\ 3 & 0 & 3 \end{bmatrix}.$$

(a)

Find the variance of $X_1 - 2X_2 + X_3$.

(b)

Find the variance matrix of $\boldsymbol{Y} = (Y_1, Y_2)^T$ where $Y_1 = X_1 + X_2$ and $Y_2 = X_1 + X_2 + X_3$.

Solution:

1. Let $\boldsymbol{a} = (1, -2, 1)^T$. Then $X_1 - 2X_2 + X_3 = \boldsymbol{a}^T \boldsymbol{X}$ and

$$Var[X_1 - 2X_2 + X_3] = Var[\boldsymbol{a}^T \boldsymbol{X}] = \boldsymbol{a}^T Var[\boldsymbol{X}]\boldsymbol{a} = 18.$$

2. Let

$$B = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

   Then $\boldsymbol{Y} = B\boldsymbol{X}$ and

$$Var[\boldsymbol{Y}] = Var[B\boldsymbol{X}] = BVar[\boldsymbol{X}]B^T = \begin{bmatrix} 12 & 15 \\ 15 & 21 \end{bmatrix}.$$