JTMS-12: Probability and Random Processes

Fall 2020

M. Bode

Lecture 18

Chapter 4: **Expectation and Introduction to Estimation**

Schwarz Inequality

Consider

$$Z = \sum_{i=1}^{n} a_i X_i$$

Interesting perspective on

$$Var[Z] = Var\left[\sum_{i=1}^{n} a_i X_i\right]$$

$$= \sum_{i,j=1}^{n} Cov[a_i X_i, a_j X_j] = \sum_{i,j=1}^{n} a_i a_j Cov[X_i, X_j]$$

$$= (a_1, \ldots, a_n) \underbrace{\begin{pmatrix} Cov[X_1, X_1] & \cdots & Cov[X_1, X_n] \\ \vdots & \ddots & \vdots \\ Cov[X_n, X_1] & \cdots & Cov[X_n, X_n] \end{pmatrix}}_{symmetric\ matrix\ M} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

Lecture 18

Chapter 4: **Expectation and Introduction to Estimation**

Schwarz Inequality

In terms of the vector $(a_1, \dots, a_n)^T$, this is a so-called quadratic form:

$$Var[Z] = (a_1, \dots, a_n) M \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

Realize: No matter what … this one (being a variance) has to be non-negative. Such forms are called positive semi-definite.

Real, symmetric matrices can always be diagonalized, and have purely real eigenvalues and eigenvectors.

Consider such an eigenvalue λ and a corresponding eigenvector v

$$0 \leq v^T M v = v^T \lambda v = \lambda \|v\|^2 \Rightarrow \lambda \geq 0$$

That is, for a positive semi-definite form, all the eigenvalues of the corresponding real symmetric matrix M have to be non-negative.

Lecture 18

Chapter 4: **Expectation and Introduction to Estimation**

Schwarz Inequality

2x2 case:

$$M = \begin{pmatrix} Cov[X_1\,,X_1] & Cov[X_1\,,X_2] \\ Cov[X_2\,,X_1] & Cov[X_2\,,X_2] \end{pmatrix}$$

Eigenvalues $\lambda_1, \lambda_2 \geq 0$

$$\Rightarrow trace(M) = \lambda_1 + \lambda_2 \geq 0, \qquad det(M) = \lambda_1 \lambda_2 \geq 0$$

Hence,

$$Cov[X_1\,,X_1]Cov[X_2\,,X_2] - Cov[X_2\,,X_1]Cov[X_1\,,X_2] \geq 0$$

or

$$\boxed{Cov^2[X_1\,,X_2] \leq Var[X_1]Var[X_2]}$$

Lecture 18

Chapter 4: **Expectation and Introduction to Estimation**

Schwarz Inequality

… and similar observations

$$E\left[\left(\sum_{i=1}^{n} a_i X_i\right)^2\right] = (a_1\,, \dots, a_n)\underbrace{\begin{pmatrix} E[X_1^2] & \cdots & E[X_1X_n] \\ \vdots & \ddots & \vdots \\ E[X_nX_1] & \cdots & E[X_2^2] \end{pmatrix}}_{symmetric\ matrix}\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

Same line of thought …

$$\boxed{E^2[X_1X_2] \leq E[X_1^2]E[X_2^2]}$$

Lecture 18

Chapter 4: **Expectation and Introduction to Estimation**

Schwarz Inequality

Interpretation:

Random variables X, Y can be seen as vectors in an (infinite dimensional) vector space.

The joint moment

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x, y) dx\, dy$$

of two random variables (if it exists) is a scalar (or inner) product, and

$$\frac{E[XY]}{\sqrt{E[X^2]E[Y^2]}} = cos(\varphi)$$

defines an angle enclosed by X and Y.

$$E[XY] = 0 \;\Leftrightarrow X \perp Y$$

for non-zero X, Y.

Lecture 18

Chapter 4: **Expectation and Introduction to Estimation**

Schwarz Inequality

Interpretation:

In this sense,

$$Cov[X, Y]$$

is the scalar (inner) product of $(X - \mu_X)$ and $(Y - \mu_Y)$

enclosing an angle φ, such that

$$cos(\varphi) = \frac{Cov[X, Y]}{\sigma_X \sigma_Y}$$

Hence, we interpret:

$$\rho = \frac{Cov[X, Y]}{\sigma_X \sigma_Y} = cos(\varphi)$$

Lecture 18

Chapter 4: **Expectation and Introduction to Estimation**

Moment Generating Functions

$$\theta_X(t) \stackrel{\text{def}}{=} E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x)dx\,, \qquad \text{for } t\epsilon\mathbb{C}$$

… usually, this exists only for some $t\epsilon\mathbb{C}$.

Mind: Except for $t \leftrightarrow -t$, this is the so-called bi-lateral Laplace transform of the pdf.

Relation to moments … if they exist:

$$E[e^{tX}] = E\left[1 + tX + \frac{1}{2}(tX)^2 + \frac{1}{6}(tX)^3 + \cdots\right]$$

$$= 1 + t\mu + \frac{t^2}{2}E[X^2] + \frac{t^3}{6}E[X^3] + \cdots$$

$$\Rightarrow E[X^k] = \frac{d^k}{dt^k}\theta_X(t)\Big|_{t=0}$$

Lecture 18

Chapter 4: **Expectation and Introduction to Estimation**

Example 1:

Consider a normal r.v. $X \sim \mathcal{N}(\mu, \sigma^2)$

$$\theta_X(t) = E[e^{tX}] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{tx} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

Complete the square as we did before …

$$\theta_X(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2 - 2\mu x + \mu^2 - 2\sigma^2 tx}{2\sigma^2}\right) dx$$

$$= \exp(\mu t + \sigma^2 t^2/2) \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}\left(x - (\mu + \sigma^2 t)\right)^2\right) dx$$

$$\Rightarrow \theta_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$$

Lecture 18

Chapter 4: **Expectation and Introduction to Estimation**

Example 1 contd …

$$\theta_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$$

Find

$$E[X^0] = \theta(0) \qquad = 1$$

$$E[X] = \theta'(0) \qquad = (\mu + \sigma^2 t)\ \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)\Bigg|_{t=0} = \mu$$

$$E[X^2] = \theta''(0) \qquad = \{\sigma^2 + (\mu + \sigma^2 t)^2\}\ \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)\Bigg|_{t=0}$$

$$= \sigma^2 + \mu^2$$

➔ $Var(X) = ?$

Lecture 18

Chapter 4: **Expectation and Introduction to Estimation**

4.1 Expected Value of a R.V.

4.2 Conditional Expectations

4.3 Moments

4.4 Chebyshev & Schwarz

4.5 Moment Generating Functions

4.6 Chernoff Bound

4.7 Characteristic Functions & Central Limit Theorem

4.8 Estimators for Mean and Variance

Example 2:

Consider a binomial r.v. X with parameters k, n, p

$$\theta_X(t) = E[e^{tX}] = \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k q^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} (e^t p)^k q^{n-k}$$

$$\Rightarrow \theta_X(t) = (e^t p + q)^n$$

$$E[X^0] = \theta(0) \ = 1$$

$$E[X] = \theta'(0) \ = n(e^t p + q)^{n-1} e^t p \Big|_{t=0} = np$$

$$E[X^2] = \theta''(0) = n(n-1)(e^t p + q)^{n-2} e^{2t} p^2 + n(e^t p + q)^{n-1} e^t p \Big|_{t=0}$$

$$= npq + n^2 p^2$$

Lecture 18

Chapter 4: **Expectation and Introduction to Estimation**

Example 3:

Consider two independent r.v.s X and Y with pdfs $f_X(x)$ and $f_Y(y)$.

Find the MGF of their sum $Z = X + Y$.

$$\theta_Z(t) = E\left[e^{t(X+Y)}\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t(x+y)} f_X(x) f_Y(y) dx\, dy$$

$$= \int_{-\infty}^{\infty} e^{tx} f_X(x) dx \int_{-\infty}^{\infty} e^{ty} f_Y(y) dy = \theta_X(t)\theta_Y(t)$$

Lecture 18

Chapter 4: **Expectation and Introduction to Estimation**

Apply …

From this perspective, interpret the previous result:

For a binomial r.v. X: $\theta_X(t) = (e^t p + q)^n$

What is the MGF of a sum of *n* i.i.d normals $X \sim \mathcal{N}(\mu, \sigma^2)$

$$\mathrm{Z} = \sum_{i=1}^{n} X_i$$

For each summand:

$$\theta_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$$

$$\theta_Z(t) = [\theta_X(t)]^n = \exp\left(n\mu t + n\frac{\sigma^2 t^2}{2}\right)$$

What‘s that?

Lecture 18

Chapter 4: **Expectation and Introduction to Estimation**

(Joint) Moment Generating Functions

$$\theta_{XY}(t_1, t_2) \stackrel{\text{def}}{=} E[\exp(t_1 X + t_2 Y)]$$

$$= \int_{-\infty}^{\infty} \exp(t_1 x + t_2 y) f_{XY}(x, y) dx dy$$

Relation to moments … if they exist:

$$E[X] = \frac{\partial}{\partial t_1} \theta_{XY}(t_1, t_2)\Big|_{(t_1,t_2)=(0,0)} \qquad E[X^2] = \frac{\partial^2}{\partial t_1^2} \theta_{XY}(t_1, t_2)\Big|_{(t_1,t_2)=(0,0)}$$

$$E[Y] = \frac{\partial}{\partial t_2} \theta_{XY}(t_1, t_2)\Big|_{(t_1,t_2)=(0,0)} \qquad E[Y^2] = \frac{\partial^2}{\partial t_2^2} \theta_{XY}(t_1, t_2)\Big|_{(t_1,t_2)=(0,0)}$$

$$E[XY] = \frac{\partial^2}{\partial t_1 \partial t_2} \theta_{XY}(t_1, t_2)\Big|_{(t_1,t_2)=(0,0)}$$

The End

Next time: Continue with Chp. 4