JTMS-12: Probability and Random Processes

Fall 2020

M. Bode

Lecture 19

Recap

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

Moment Generating Functions

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

Relation to moments … if they exist:

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

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

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

Lecture 19

Recap

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

(Joint) Moment Generating Functions

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

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

Relation to moments … if they exist:

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

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

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

Lecture 19

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

Chernoff Bound

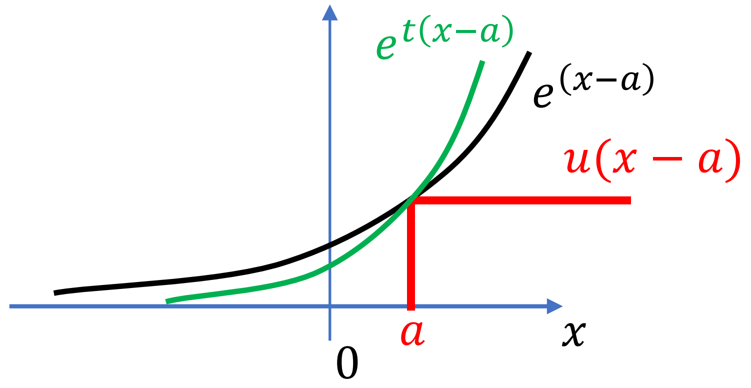
Another bound for the small tails probability …

$$P[X \geq a] = \int_a^\infty f_X(x)dx$$

$$= \int_{-\infty}^{\infty} u(x-a) f_X(x)dx \leq \int_{-\infty}^{\infty} e^{(x-a)} f_X(x)dx$$

Also,

$$P[X \geq a] \leq \int_{-\infty}^{\infty} e^{t(x-a)} f_X(x)dx\,, \qquad \text{for } t \geq 0$$



Lecture 19

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

Chernoff Bound

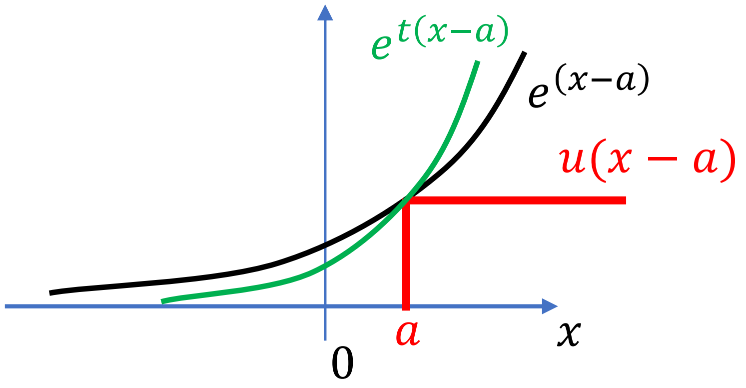
Another bound for the small tails probability …

Hence, for $t \geq 0$,

$$\boxed{P[X \geq a] \leq e^{-at} \int_{-\infty}^{\infty} e^{tx} f_X(x)dx = e^{-at}\theta_X(t)}$$

Discrete version:

$$\boxed{P[X \geq a] \leq e^{-at} \sum_{k=0}^{\infty} e^{tx_k} P[X = x_k] = e^{-at}\theta_X(t)}$$



Lecture 19

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

Chernoff Bound(s)

For $t \geq 0$,

$$\boxed{P[X \geq a] \leq e^{-at} \theta_X(t)}$$

Notice:

This inequality, actually, offers infinitely many different (*upper*) bounds.

Which one is most relevant/interesting?

The lowest one … yes!

Final task … find the lowest Chernoff bound!

Lecture 19

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

Chernoff Bound(s)

Example: Consider $X \sim \mathcal{N}(\mu, \sigma^2)$

In the previous lecture, we found

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

For $t \geq 0$,

$$\boxed{P[X \geq a] \leq e^{-at}\theta_X(t)}$$

So, the family of Chernoff bounds is

$$P[X \geq a] \leq e^{-at}\theta_X(t) = \exp\left((\mu - a)t + \frac{\sigma^2 t^2}{2}\right)$$

Lecture 19

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

Chernoff Bound(s)

Example: Consider $X \sim \mathcal{N}(\mu, \sigma^2)$

… the family of Chernoff bounds is

$$P[X \geq a] \leq e^{-at}\theta_X(t) = \exp\left((\mu - a)t + \frac{\sigma^2 t^2}{2}\right), \qquad t \geq 0$$

Which value of t helps most?

Derive with respect to t ?

$$\frac{\partial}{\partial t}\exp\left((\mu - a)t + \frac{\sigma^2 t^2}{2}\right)$$

$$= \left((\mu - a) + t\sigma^2\right)\exp\left((\mu - a)t + \frac{\sigma^2 t^2}{2}\right) = 0$$

Lecture 19

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

Chernoff Bound(s)

Example: Consider $X \sim \mathcal{N}(\mu, \sigma^2)$

$$= \left((\mu - a) + t\sigma^2\right) \exp\left((\mu - a)t + \frac{\sigma^2 t^2}{2}\right) = 0$$

$$\Leftrightarrow (\mu - a) + t\sigma^2 = 0$$

This leads to

$$t_{opt} = \frac{a - \mu}{\sigma^2}$$

… and a corresponding bound

$$\exp\left((\mu - a)t_{opt} + \frac{\sigma^2 t_{opt}^2}{2}\right) = \exp\left(-\frac{(a - \mu)^2}{\sigma^2} + \frac{(a - \mu)^2}{2\sigma^2}\right)$$

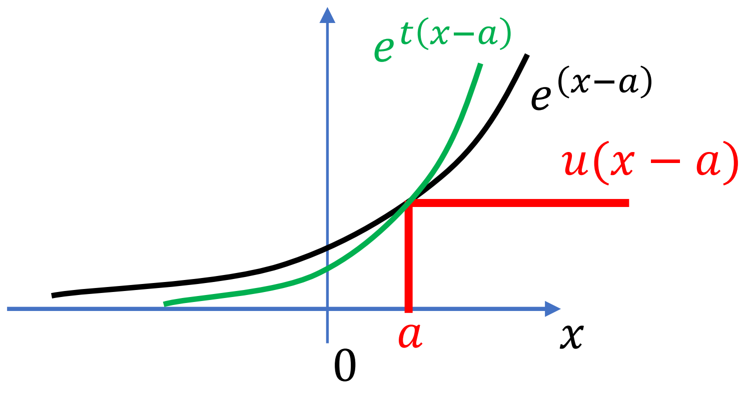
$$= \exp\left(-\frac{(a - \mu)^2}{2\sigma^2}\right)$$

Lecture 19

Recall …

$$P[X \geq a] = \int_a^\infty f_X(x)dx$$

$$= \int_{-\infty}^{\infty} u(x-a) f_X(x)dx \leq \int_{-\infty}^{\infty} e^{(x-a)} f_X(x)dx$$



… the family of Chernoff bounds is ($t \geq 0$)

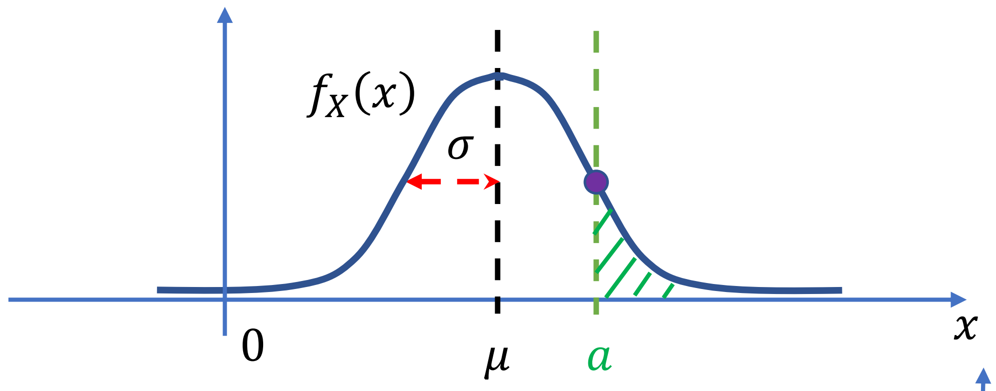
$$P[X \geq a] \leq e^{-at}\theta_X(t) = \exp\left((\mu - a)t + \frac{\sigma^2 t^2}{2}\right)$$

Chernoff Bound(s)

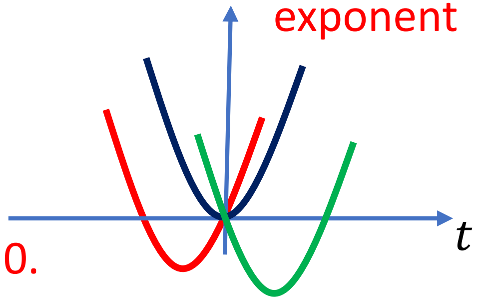
Example: Consider $X \sim \mathcal{N}(\mu, \sigma^2)$

… best Chernoff bound:

$$\boldsymbol{P[X \geq a]} \leq \exp\left(-\frac{(a-\mu)^2}{2\sigma^2}\right) = \sqrt{2\pi\sigma^2}\,\boldsymbol{f_X(x=a)}$$



Discuss exponent



Does that make sense?

Looks OK for $a > \mu$ … but what if $a < \mu$?

… t negative … replace by boundary point at t = 0.

Lecture 19

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

Characteristic Functions

$$\Phi_X(\omega) \stackrel{\text{def}}{=} E\left[e^{j\omega X}\right] = \int_{-\infty}^{\infty} e^{j\omega x} f_X(x)dx$$

if the integral exists.

Mind: Except for the sign of ω, this is the so-called Fourier transform of the pdf.

Relation to moments … if they exist:

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

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

$$\Rightarrow E[X^k] = \frac{1}{j^k}\frac{d^k}{d\omega^k}\Phi_X(\omega)\bigg|_{t=0}$$

Lecture 19

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

Example:

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

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

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

$$= \int_{-\infty}^{\infty} e^{j\omega x} f_X(x) dx \int_{-\infty}^{\infty} e^{j\omega y} f_Y(y) dy = \Phi_X(\omega)\Phi_Y(\omega)$$

➔ Convolution of pdfs // product of characteristic functions

Lecture 19

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

Recall…

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

Apply …

What is the characteristic function of a sum of *n* i.i.d standard normals $X \sim \mathcal{N}(0,1)$

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

For each summand:

$$\Phi_X(\omega) = \exp\left(-\frac{\omega^2}{2}\right)$$

$$\Rightarrow \Phi_Z(t) = [\Phi_X(t)]^n = \exp\left(-n\frac{\omega^2}{2}\right)$$

What's that?

… the characteristic function of a Gaussian with $\mu_Z = 0$, and $\sigma_Z^2 = n$

$$f_Z(z) = \frac{1}{\sqrt{2\pi n}} \exp\left(-\frac{z^2}{2n}\right)$$

Lecture 19

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

(Joint) Caracteristic Functions

$$\Phi_{XY}(\omega_1, \omega_2) \overset{\text{def}}{=} E[\exp(j\omega_1 X + j\omega_2 Y)]$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(j\omega_1 x + j\omega_2 y) f_{XY}(x, y) dx\, dy$$

Relation to moments … if they exist:

$$E[X] = \frac{1}{j} \frac{\partial}{\partial \omega_1} \Phi_{XY}(\omega_1, \omega_2)\Big|_{(\omega_1,\omega_2)=(0,0)} \quad E[X^2] = -\frac{\partial^2}{\partial \omega_1^2} \Phi_{XY}(\omega_1, \omega_2)\Big|_{(\omega_1,\omega_2)=(0,0)}$$

$$E[Y] = \frac{1}{j} \frac{\partial}{\partial \omega_2} \Phi_{XY}(\omega_1, \omega_2)\Big|_{(\omega_1,\omega_2)=(0,0)} \quad E[Y^2] = -\frac{\partial^2}{\partial \omega_2^2} \Phi_{XY}(\omega_1, \omega_2)\Big|_{(\omega_1,\omega_2)=(0,0)}$$

$$E[XY] = -\frac{\partial^2}{\partial \omega_1 \partial \omega_2} \Phi_{XY}(\omega_1, \omega_2)\Big|_{(\omega_1,\omega_2)=(0,0)}$$

Lecture 19

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

Example:

Consider two jointly (zero-mean) Gaussian r.v.s V and W with a joint pdf like, e.g.,

$$f_{VW}(v,w) = \frac{1}{8\pi}\exp\left[-\frac{5v^2+6vw+5w^2}{32}\right]$$

Find their joint characteristic function $\Phi_{VW}(\omega_1,\omega_2)$

$$\Phi_{VW}(\omega_1,\omega_2) \stackrel{\text{def}}{=} E[\exp(j\omega_1 V + j\omega_2 W)]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp(j\omega_1 v + j\omega_2 w) f_{VW}(v,w)\,dv\,dw$$

… can be done by completing the squares … twice.

But we try a more elegant *and telling* approach …

Lecture 19

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

Recall...

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

Example … contd:

Let's introduce

$$Z = \omega_1 V + \omega_2 W$$

Realize …

$$\Phi_{VW}(\omega_1, \omega_2) \stackrel{\text{def}}{=} E[\exp(j\omega_1 V + j\omega_2 W)]$$

$$= E[\exp(jZ)] = E[\exp(j\omega Z)]\Big|_{\omega=1} \stackrel{\text{def}}{=} \Phi_Z(\omega)\Big|_{\omega=1}$$

As Z is Gaussian and zero-mean,

$$\Phi_Z(\omega) = \exp\left(-\frac{\sigma_Z^2 \omega^2}{2}\right)$$

Now,

$$\sigma_Z^2 = \mathrm{Var}[\omega_1 V + \omega_2 W]$$

$$= \omega_1^2 \mathrm{Var}[V] + 2\omega_1\omega_2 Cov[V, W] + \omega_2^2 \mathrm{Var}[W]$$

Lecture 19

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

Example … contd:

Hence,

$$\Phi_{VW}(\omega_1,\omega_2) = \Phi_Z(\omega)\Big|_{\omega=1} = \exp\left(-\frac{\sigma_Z^2}{2}\right)$$

$$= \exp\left(-\frac{1}{2}\{\omega_1^2\mathrm{Var}[V] + 2\omega_1\omega_2 Cov[V,W] + \omega_2^2\mathrm{Var}[W]\}\right)$$

$$= \exp\left(-\frac{1}{2}\left\{(\omega_1,\omega_2)\begin{pmatrix}\mathrm{Var}[V] & Cov[V,W]\\ Cov[V,W] & \mathrm{Var}[W]\end{pmatrix}\begin{pmatrix}\omega_1\\ \omega_2\end{pmatrix}\right\}\right)$$

How large are $\mathrm{Var}[V]$, $\mathrm{Var}[W]$, $Cov[V,W]$?

Lecture 19

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

Example … contd:

$$\Phi_{VW}(\omega_1,\omega_2) = \exp\left(-\frac{1}{2}\left\{(\omega_1,\omega_2)\begin{pmatrix}\mathrm{Var}[V] & Cov[V,W] \\ Cov[V,W] & \mathrm{Var}[W]\end{pmatrix}\begin{pmatrix}\omega_1 \\ \omega_2\end{pmatrix}\right\}\right)$$

How large are $\mathrm{Var}[V]$, $\mathrm{Var}[W]$, $Cov[V,W]$?

Recall, our example resulted from **task 9 in lecture 13**:

Consider two independent standard normal r.v.s X, Y

$$f_{XY}(x,y) = \frac{1}{2\pi}\exp\left[-\frac{x^2+y^2}{2}\right]$$

$$V = g(X,Y) = X + 2Y, \qquad W = h(X,Y) = X - 2Y$$

Hence, $\mathrm{Var}[V] = 5$, $\mathrm{Var}[W] = 5$, $Cov[V,W] = -3$

$$\Rightarrow \Phi_{VW}(\omega_1,\omega_2) = \exp\left(-\frac{1}{2}\{5\omega_1^2 - 6\omega_1\omega_2 + 5\omega_2^2\}\right)$$

Lecture 19

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

Example … contd:

Compare:

$$f_{VW}(v,w) = \frac{1}{8\pi}\exp\left[-\frac{5v^2+6vw+5w^2}{2\cdot 16}\right]$$

$$\Phi_{VW}(\omega_1,\omega_2) = \exp\left(-\frac{1}{2}\{5\omega_1^2 - 6\omega_1\omega_2 + 5\omega_2^2\}\right)$$

The similarity is not accidental.

… we‘ll see the general relation in chapter 5 on random vectors.

Any guesses so far?

The End

Next time: Continue with Chp. 4