S722 HW8

John Koo

To save on typing, I will denote $\frac{\partial^k}{\partial x^k} f(x) = \partial_x^k f(x)$.

3.45

a

$M_X(t) = \int e^{tx} f(x) dx \geq \int_a^\infty e^{tx} f(x) dx \geq e^{at} \int_a^\infty f(x) dx = e^{at} P(X \geq a)$
$\implies e^{-at} M_X(t) \geq P(X \geq a)$

b

$M_X(t) = \int e^{tx} f(x) dx \geq \int_{-\infty}^a e^{tx} f(x) dx \geq e^{ta} P(X \leq a)$
$\implies e^{-at} M_X(t) \geq P(X \leq a)$

Example 5.5.8

Convergence in probability

Let $\epsilon > 0$ and $\delta \in (0, 1]$.

$|X_n - X| = X_n - X$ is either 0 or 1, and the probability that it is 1 follows this pattern:

$$\begin{array}{cccc} 1 & & & \\ 1/2 & 1/2 & & \\ 1/3 & 1/3 & 1/3 & \\ 1/4 & 1/4 & 1/4 & 1/4 \\ \vdots & & & \end{array}$$

So $X_n - X = 1$ with probability $1/k$ where $\frac{k(k+1)}{2} = n \implies k = \lceil \frac{-1+\sqrt{1+8n}}{2} \rceil$, and $X_n - X = 0$ with probability $1 - 1/k$.

Letting $k = \lceil \frac{1}{\delta} \rceil$ and $N = \frac{k(k+1)}{2}$, we get that $\forall n > N$, $P(|X_n - X| < \epsilon) > 1 - \delta$.

Nonconvergence almost surely

Let $N \in \mathbb{N}$ and $\epsilon \in (0, 1)$. Then we can find some $n > N$ such that $X_n - X = 1 > \epsilon$. Therefore, X_n does not converge pointwise to $X \implies X_n$ does not converge almost surely to X.

5.39

a

h is continuous $\implies \forall \epsilon > 0, \exists \delta$ s.t. $|x_n - x| < \delta \implies |h(x_n) - h(x)| < \epsilon$.

Then if $P(|X_n - X| < \delta) \to 1$, $P(|h(X_n) - h(X)| < \epsilon) \to 1$.

Therefore, $h(X_n) \xrightarrow{p} h(X)$.

b

Suppose instead, we have the sequence $X_n(s) = s + I(s < \frac{1}{n})$, which we can see is a subsequence of the one in example 5.5.8. Then $X_n(s)$ converges to $X(s)$ pointwise for $s \neq 0$, so $X_n \overset{a.s.}{\to} X$.

To prove pointwise convergence, suppose $\epsilon > 0$. Then let $N = 1/\epsilon$. For any $n > N$, $X_n(s) = s$ when $s > \epsilon$.

7.41

a

$E[\sum_i a_i X_i] = \sum a_i E[X_i] = \mu \sum a_i = \mu$

b

We can see that $Var(\sum a_i X_i) = \sum a_i^2 Var(X_i) = \sigma^2 \sum a_i^2$ so the objective is to minimize $\sum a_i^2$ with the constraint $\sum a_i = 1$.

This can be solved by setting $\nabla(\sum a_i^2 + \lambda(\sum a_i - 1)) = 0 \implies \begin{bmatrix} 2a_1 - \lambda \\ \vdots \\ 2a_n - \lambda \end{bmatrix} = \vec{0}$.

The sum of these elements yields 0, so we get:

$0 = \sum(2a_i - \lambda)$
$= 2\sum a_i - n\lambda$
$= 2 - n\lambda$
$\implies \lambda = 2/n$

Plugging this into any one of the elements, we get that $a_i = 1/n$. So $\frac{\sum X_i}{n}$ has the lowest variance.

7.42

a

Similar to the previous problem, we want to minimize

$Var(\sum a_i W_i) = \sum a_i^2 \sigma_i^2$,

under the constraint

$\sum a_i = 1$.

Then the equation we need to solve is

$\nabla(\sum a_i^2 \sigma_i^2 - \lambda(\sum a_i - 1)) = 0$
$\implies 2a_i\sigma_i^2 - \lambda = 0 \; i = 1, ..., n$
$\implies a_i - \frac{\lambda}{2\sigma_i^2} = 0$

$\implies \sum a_i - \frac{\lambda}{2}\sum \frac{1}{\sigma_i^2} = 0$
$\implies \lambda = \frac{2}{\sum 1/\sigma_i^2}$

Plugging this back into $2a_i\sigma_i^2 - \lambda = 0$, we get
$2a_i\sigma_i^2 - 2/(\sum_j 1/\sigma_j^2) = 0$
$\implies a_i = \frac{1/\sigma_i^2}{\sum_j 1/\sigma_j^2}$

b

$Var(\frac{\sum W_i/\sigma_i^2}{\sum 1/\sigma_i^2})$
$= (\frac{1}{\sum 1/\sigma_i^2})^2 \sum Var(W_i/\sigma_i^2)$
$= (\frac{1}{\sum 1/\sigma_i^2})^2 \sum (1/\sigma_i^2)^2\sigma_i^2$
$= (\frac{1}{\sum 1/\sigma_i^2})^2 \sum 1/\sigma_i^2$
$= \frac{1}{\sum 1/\sigma_i^2}$

$\rho = 0 \implies$ independence for bivariate normal

$f_{X,Y}(x,y) \propto \exp\Bigg(-\frac{1}{2(1-\rho^2)}\big((\frac{x-\mu_X}{\sigma_X})^2 - 2\rho\frac{x-\mu_X}{\sigma_X}\frac{y-\mu_Y}{\sigma_Y} + (\frac{y-\mu_Y}{\sigma_Y})^2\big)\Bigg)$

Setting $\rho = 0$, we can clearly see that the cross term goes away, so we get

$f_{X,Y}(x,y) \propto e^{-\frac{1}{2}(z_X^2+z_Y^2)} = e^{-z_X^2/2}e^{-z_Y^2/2}$ where $z_X = \frac{x-\mu_X}{\sigma_X}$ and z_Y defined similarly for y. This is separable into $g(x)$ and $h(y)$, so X and Y are independent.

Theorem 5.5.2

Let $X_1, ..., X_n \overset{iid}{\sim} F(x)$ with $E[X_i] = \mu$ and $Var(X_i) = \sigma^2 < \infty$. Then $\bar{X}_n \xrightarrow{p} \mu$.

proof

Let $\epsilon > 0$.

Then $P(|\bar{X}_n - \mu| \geq \epsilon) = P((\bar{X}_n - \mu)^2 \geq \epsilon^2) \leq \frac{E[(\bar{X}_n-\mu)^2]}{\epsilon^2} = \frac{Var(\bar{X}_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}$

Let $\delta > 0$ and $N = \frac{\sigma^2}{\delta\epsilon^2}$. Then $\forall n > N$, $P(|\bar{X}_n - \mu| \geq \epsilon) < \delta$.

5.24

Letting $U = X_{(1)}$ and $V = X_{(n)}$, we are given that $f_{UV}(u,v) = \frac{n(n-1)}{\theta^n}(v-u)^{n-2}I(u<v)$.

Let $T = U/V$ and $S = V$. Then $V = S$ and $U = TS$, so

$\partial_T U = S$
$\partial_S U = T$
$\partial_T V = 0$
$\partial_S V = 1$
$\implies |J| = s$

Then $f_{TS}(t,s) \propto (s - ts)^{n-2}I(ts < s)s = s^{n-1} \times (1-t)^{n-2}I(t<1)$, which is separable into f_T and f_S.

5.25

Let $Y_i = X_{(i)}$ and $Y = (Y_1, ..., Y_n)$.

Then $f_Y(y_1, ..., y_n) = n! a^n \theta^{-an} \prod_i y_i^{a-1}$

Let $Z_i = Y_i / Y_{i+1}$ for $i \leq n-1$ and $Z_n = Y_n$ and $Z = (Z_1, ..., Z_n)$..
Then $Y_n = Z_n$, $Y_{n-1} = Z_{n-1} Z_n$, $Y_{n-2} = Z_{n-2} Z_{n-1} Z_n$, etc.

Then $|J| = z_2 z_3^2 \cdots z_n^{n-1}$.

So $f_Z(z_1, ..., z_n) \propto (z_1 \cdots z_n)^{a-1} (z_2 \cdots z_n)^{a-1} \cdots z_n^{a-1} z_2 z_3^2 \cdots z_n^{n-1}. = z_1^{p_1} z_2^{p_2} \cdots z_n^{p_n}$ for some $p_1, ..., p_n$, which is separable by each z_i.

8.5

c

We saw in the previous part that T can be written as

$T = \sum_i \log X_i - n \log X_{(1)}$
$= \sum_i (\log X_i - \log X_{(1)})$

Let $Y = \log X_i$.
Then $X = e^Y$.
and $X' = e^Y$.
Then $f_Y(y) = \nu e^{-y} I(y \geq \nu)$.

Now let $Z_i = Y_i - Y_{(1)}$.
Then the indicator function goes away, so we have $f_Z(z) = e^{-z}$, i.e., $Z_i \sim \mathit{Exponential}(1)$. Therefore, $T \sim \mathit{Gamma}(n-1, 1)$, and $2T \sim \mathit{Gamma}(n-1, 2) = \chi^2_{2(n-1)}$.