MATH 507a Homework 9

Qilin Ye

December 5, 2022

Problem: Durrett 3.10.5

Let φ be the ch.f. of a distribution F on $\mathbb{R}$. What is the distribution on $\mathbb{R}^d$ that corresponds to the ch.f. $\psi(t_1, ..., t_d) = \varphi(t_1 + ... + t_d)$?

Solution. The distribution $(X, X, ..., X)$ on $\mathbb{R}^d$ has ch.f.

$$\psi(t_1, ..., t_d) = \exp(i1^T(t_1, ..., t_d)) = \exp(i(t_1 + ... + t_d)) = \varphi(t_1 + ... + t_d).$$

Problem: Durrett 4.1.2

Prove Chebyshev's inequality: if $a > 0$ then

$$\mathbb{P}(|X| \geqslant a \mid \mathcal{F}) \leqslant a^{-2}\mathbb{E}(X^2 \mid \mathcal{F}).$$

Proof. By definition $\mathbb{P}(|X| \geqslant a \mid \mathcal{F}) = \mathbb{E}(1_{|X|\geqslant a} \mid \mathcal{F})$. For any $F \in \mathcal{F}$, the normal Chebyshev's inequality gives

$$\int_F 1_{|X|\geqslant a} \, \mathrm{d}\mathbb{P} \leqslant \int_F a^{-2}X^2 \, \mathrm{d}\mathbb{P}.$$

Therefore $\mathbb{P}(|X| \geqslant a \mid \mathcal{F}) = \mathbb{E}(1_{|X|\geqslant a} \mid \mathcal{F}) \leqslant a^{-2}\mathbb{E}(X^2 \mid \mathcal{F})$. □

Problem: Durrett 4.1.8

Let $Y_1, Y_2, ...$ be i.i.d. with mean μ and variance σ^2. Let N be an independent positive integer valued r.v. with $\mathbb{E}N^2 < \infty$ and $X = Y_1 + ... + Y_N$. Show that $\mathrm{var}(X) = \sigma^2\mathbb{E}N + \mu^2 \mathrm{var}(N)$.

Proof. Since $\mathbb{E}(X \mid N = n) = n\mu$ and $\mathrm{var}(X \mid N = n) = n\sigma^2$, we obtain $\mathbb{E}(X \mid N) = N\mu$ and $\mathrm{var}(X \mid N) = N\sigma^2$. Using the hint,

$$\mathrm{var}(X) = \mathbb{E}(\mathrm{var}(X \mid \mathcal{F})) + \mathrm{var}(\mathbb{E}(X \mid \mathcal{F})) = \sigma^2\mathbb{E}N + \mu^2 \mathrm{var}(N).$$ □

Problem 1

Suppose $X^{(n)} \sim \mathcal{N}(\mu^{(n)}, \Sigma^{(n)})$ in $\mathbb{R}^d$ and for some $X^{(\infty)}$ we have $\mathbb{E}[|X^{(n)} - X^{(\infty)}|^2] \to 0$. Let $\mu^{(\infty)}$, $\Sigma^{(\infty)}$ be the mean and covariance matrix of $X^{(\infty)}$.

(1) Show that $\mu^{(n)} \to \mu^{(\infty)}$ and $\Sigma^{(n)} \to \Sigma^{(\infty)}$.

(2) Show that $X^{(\infty)}$ is multivariate normal.

Proof. (1) We have $\mathbb{E}|X^{(\infty)}| < \infty$ and in particular each component $\mathbb{E}|X_i^{(\infty)}| < \infty$. By Hölder's inequality, $\mathbb{E}|X^{(n)} - X^{(\infty)}|^2 \to 0$ implies $\mathbb{E}|X^{(n)} - X^{(\infty)}| \to 0$ as well. In particular, each component $\mathbb{E}|X_i^{(n)} - X_i^{(\infty)}| \to 0$ and is bounded in n. This establishes $\mu^{(n)} \to \mu^{(\infty)}$. Next up,

$$\begin{aligned}\mathbb{E}(X_i^{(n)} X_j^{(n)}) &= \mathbb{E}[(X_i^{(\infty)} + (X_i^{(n)} - X_i^{(\infty)}))(X_j^{(\infty)} + (X_j^{(n)} - X_j^{(\infty)}))] \\ &= \mathbb{E}(X_i^{(\infty)} X_j^{(\infty)}) + \mathbb{E}X_i^{(\infty)}(X_j^{(n)} - X_j^{(\infty)}) \\ &\quad + \mathbb{E}X_j^{(\infty)}(X_i^{(n)} - X_i^{(\infty)}) + \mathbb{E}(X_i^{(n)} - X_i^{(\infty)})(X_j^{(n)} - X_j^{(\infty)}).\end{aligned}$$

By Cauchy-Schwarz

$$\mathbb{E}X_i^{(\infty)}(X_j^{(n)} - X_j^{(\infty)}) \leqslant \sqrt{\mathbb{E}|X_i^{(\infty)}|^2}\sqrt{\mathbb{E}|X_j^{(n)} - X_j^{(\infty)}|^2} \to 0,$$

and similarly

$$\mathbb{E}(X_i^{(n)} - X_i^{(\infty)})(X_j^{(n)} - X_j^{(\infty)}) \leqslant \sqrt{\mathbb{E}|X_i^{(n)} - X_i^{(\infty)}|^2}\sqrt{\mathbb{E}|X_j^{(n)} - X_j^{(\infty)}|^2} \to 0.$$

Therefore we obtain $\mathbb{E}(X_i^{(n)} X_j^{(n)}) \to \mathbb{E}(X_i^{(\infty)} X_j^{\infty})$. Finally,

$$\begin{aligned}\operatorname{cov}(X_i^{(n)}, X_j^{(n)}) &= \mathbb{E}[(X_i^{(n)} - \mathbb{E}X_i^{(n)})(X_j^{(n)} - \mathbb{E}X_j^{(n)})] \\ &= \mathbb{E}X_i^{(n)} X_j^{(n)} - \mathbb{E}X_i^{(n)}\mathbb{E}X_j^{(n)} \to \mathbb{E}X_i^{(\infty)} X_j^{(\infty)} - \mathbb{E}X_i^{(\infty)}\mathbb{E}X_j^{(\infty)},\end{aligned}$$

completing the proof of $\Sigma^{(n)} \to \Sigma^{(\infty)}$.

(2) Convergence in L^2 implies convergence in probability, which implies convergence in distribution, which then implies pointwise convergence of the characteristic functions. Since $\mu^{(n)} \to \mu^{\infty}$, $\Sigma^{(n)} \to \Sigma^{(\infty)}$, and $\exp(\cdot)$ is continuous, the limiting ch.f. must be that of a multivariate normal with parameters $\mu^{(\infty)}$ and $\Sigma^{(\infty)}$.

□

Problem 2

Suppose the characteristic function φ of the random vector $X \in \mathbb{R}^d$ satisfies $\int_{\mathbb{R}^d} |\varphi(t)| \, \mathrm{d}t < \infty$. Show that X has bounded continuous density given by

$$f(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-it^T x} \varphi(t) \, \mathrm{d}t.$$

Proof. From Durrett's theorem 3.10.4, if $A = \prod[a_i, b_i]$ with $\mu(\partial A) = 0$ then

$$\mu(A) = \lim_{T\to\infty} \frac{1}{(2\pi)^d} \int_{[-T,T]^d} \prod_{j=1}^{d} \frac{\exp(-it_j a_j) - \exp(-it_j b_j)}{it_j} \cdot \varphi(t)\, \mathrm{d}t.$$

We imitate the proof of theorem 3.3.14. Since

$$\prod_{j=1}^{d} \left| \frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j} \right| \leqslant \prod_{j=1}^{d} |b_j - a_j|,$$

the above integral converges absolutely and

$$\mu(A) + \mu(\partial A)/2 = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \prod_{j=1}^{\infty} \frac{e^{it_j a_j} - e^{it_j b_j}}{it_j} \cdot \varphi(t)\, \mathrm{d}t.$$

This implies

$$\begin{aligned} \mu(\prod_{j=1}^{d}(x_i, x_i + h_i)) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \prod_{i=1}^{d} \frac{e^{-it_i x_i} - e^{-it_i(x_i+h_i)}}{it_i} \cdot \varphi(t)\, \mathrm{d}t \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \prod_{i=1}^{d} \int_{x_i}^{x_i+h_i} e^{-it_i y_i}\, \mathrm{d}y_i \varphi(t)\, \mathrm{d}t \\ &= \int_{x}^{x+h} \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-it^T y} \varphi(t)\, \mathrm{d}y\, \mathrm{d}x_1 \ldots \mathrm{d}x_d, \end{aligned}$$

and so X has density $\frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-it^T x} \varphi(t)\, \mathrm{d}t$, as claimed. □

Problem 3

For a σ-field $\mathcal{F}$ and r.v.'s X with $\mathbb{E}X^2 < \infty$, $\mathbb{E}Y^2 < \infty$, let

$$\mathrm{var}(X \mid \mathcal{F}) = \mathbb{E}([X - \mathbb{E}(X \mid \mathcal{F})]^2 \mid \mathcal{F}) = \mathbb{E}(X^2 \mid \mathcal{F}) - (\mathbb{E}(X \mid \mathcal{F}))^2$$

and

$$\mathrm{cov}(X, Y \mid \mathcal{F}) = \mathbb{E}([X - \mathbb{E}(X \mid \mathcal{F})][Y - \mathbb{E}(Y \mid \mathcal{F})] \mid \mathcal{F}) = \mathbb{E}(XY \mid \mathcal{F}) - \mathbb{E}(X \mid \mathcal{F})\mathbb{E}(Y \mid \mathcal{F}).$$

(1) Let $\mathcal{G}_1 \subset \mathcal{G}_2$ be σ-fields. Show that $\mathbb{E}[\mathrm{var}(X \mid \mathcal{G}_1)] \geqslant \mathbb{E}[\mathrm{var}(X \mid \mathcal{G}_2)]$ almost surely. "Less information means more variability."

(2) Show that $\mathrm{cov}(X, Y) = \mathbb{E}[\mathrm{cov}(X, Y \mid \mathcal{F}) + \mathrm{cov}(\mathbb{E}(X \mid \mathcal{F}), \mathbb{E}(Y \mid \mathcal{F}))$ almost surely.

(3) If $\mathbb{E}[\mathbb{E}(X \mid \mathcal{F})^2] = \mathbb{E}X^2 < \infty$, show that $X = \mathbb{E}(X \mid \mathcal{F})$ a.s.

Proof. (1) Let $U = X - E(X \mid \mathcal{G}_2)$ and $V = \mathbb{E}(X \mid \mathcal{G}_2) - \mathbb{E}(X \mid \mathcal{G}_1)$. Since $\mathcal{G}_1 \subset \mathcal{G}_2$, V is $\mathcal{G}_2$-measurable. Then by pulling out V,

$$\mathbb{E}UV = \mathbb{E}[(\mathbb{E}(UV \mid \mathcal{G}_2)] = \mathbb{E}[V\mathbb{E}(U \mid \mathcal{G}_2)] = \mathbb{E}[0] = 0.$$

Therefore $\mathbb{E}(U + V)^2 = \mathbb{E}U^2 + \mathbb{E}V^2$, and in particular

$$\mathbb{E}(X - \mathbb{E}(X \mid \mathcal{G}_1))^2 \geqslant \mathbb{E}(X - \mathbb{E}(X \mid \mathcal{G}_2))^2.$$

This implies $\mathrm{var}(X \mid \mathcal{G}_1) \geqslant \mathrm{var}(X \mid \mathcal{G}_2)$ and taking expectation gives the claim.

(2) By definition we have

$$\operatorname{cov}(X, Y) = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y,$$

and

$$\operatorname{cov}(\mathbb{E}(X \mid \mathcal{F}), \mathbb{E}(Y \mid \mathcal{F})) = \mathbb{E}[\mathbb{E}(X \mid \mathcal{F})\mathbb{E}(Y \mid \mathcal{F})] - \mathbb{E}X\mathbb{E}Y,$$

so

$$\operatorname{cov}(X, Y) = \operatorname{cov}(\mathbb{E}(X \mid \mathcal{F}), \mathbb{E}(Y \mid \mathcal{F})) + (\mathbb{E}XY - \mathbb{E}(\mathbb{E}[X \mid \mathcal{F})\mathbb{E}(Y \mid \mathcal{F})]),$$

but the second term is precisely $\mathbb{E}(\operatorname{cov}(X, Y \mid \mathcal{F}))$ by expanding the terms and taking expectation.

(3) Define $Y = \mathbb{E}(X \mid \mathcal{F})$. Then

$$\mathbb{E}XY = \mathbb{E}[\mathbb{E}(XY \mid \mathcal{F})] = \mathbb{E}[Y\mathbb{E}(X \mid \mathcal{F})] = \mathbb{E}Y^2.$$

Therefore

$$\mathbb{E}(Y - X)^2 = \mathbb{E}Y^2 - 2\mathbb{E}XY + \mathbb{E}X^2 = \mathbb{E}X^2 - \mathbb{E}Y^2 = 0,$$

showing $Y = X$ a.s. □

Problem 4

Suppose X is a r.v. on $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{E}|X|^p < \infty$ for some $p > 0$.

(1) Show that for every event A we have $\mathbb{E}(|X|^p 1_A) = \int_0^\infty px^{p-1}\mathbb{P}(A \cap \{|X| > x\})\, \mathrm{d}x$.

(2) Show that for every σ-field $\mathcal{G} \subset \mathcal{F}$ we have $\mathbb{E}(|X|^p \mid \mathcal{G}) = \int_0^\infty px^{p-1}\mathbb{P}(|X| > x \mid \mathcal{G})\, \mathrm{d}\mathbb{P}$.

Proof. (1) This follows from Tonelli's theorem and change of variable formula:

$$\begin{aligned}\int_0^\infty px^{p-1}\mathbb{P}(A \cap \{|X| > x)\}\, \mathrm{d}x &= \int_0^\infty \int_\Omega 1_{|X|^p > x, x \in A}\, \mathrm{d}\mathbb{P}\, \mathrm{d}x \\ &= \int_\Omega \int_{-\infty}^\infty 1_{|X|^p > x, x \in A}\, \mathrm{d}x\, \mathrm{d}\mathbb{P} \\ &= \int_\Omega |X|^p 1_A\, \mathrm{d}\mathbb{P} = \mathbb{E}(|X|^p 1_A).\end{aligned}$$

(2) Let $Y = \mathbb{E}(|X|^p \mid \mathcal{G})$. Let $Z_x = \mathbb{P}(\{|X| > x \mid \mathcal{G}\}$. Then $\mathbb{P}(A \cap \{|X| > x\}) = \mathbb{E}Z_x 1_A$, so

$$\mathbb{E}(|X|^p 1_A) = \int_0^\infty px^{p-1}\mathbb{E}[Z_x 1_A]\, \mathrm{d}x.$$

For any $A \in \mathcal{G}$, we have

$$\int_A |X|^p\, \mathrm{d}\mathbb{P} = \mathbb{E}(|X|^p 1_A) = \int_A \int_0^\infty px^{p-1} Z_x\, \mathrm{d}x\, \mathrm{d}\mathbb{P},$$

so by Fubini and (1) we are done.

□

Problem 5

Let Σ be a $d \times d$ symmetric matrix and let $\lambda_1 \geqslant \ldots \geqslant \lambda_d$ be its eigenvalues, with corresponding eigenvectors $\theta^{(1)}, \ldots, \theta^{(d)}$.

(1) Show that if $\lambda_i \neq \lambda_j$ then $\theta^{(i)}, \theta^{(j)}$ are orthogonal.

(2) Suppose Σ is also positive definite. what vectors maximize $\theta^T \Sigma \theta$ over all unit vectors?

(3) Suppose Σ is the PD covariance matrix of some r.v. X, Σ is not a multiple of I, and $\theta_{\max}, \theta_{\min}$ are vectors that maximize and minimize $\operatorname{var}(\theta^T X)$ over all unit vectors θ. Show that they are orthogonal.

Proof. (1) Let $\lambda_i \neq \lambda_j$ be given. If a linear combination $c_i\theta^{(i)} + c_j\theta^{(j)} = 0$, then

$$0 = \Sigma \cdot 0 = \Sigma(c_i\theta^{(i)} + c_j\theta^{(j)}) = c_i\lambda_i\theta^{(i)} + c_j\lambda_j\theta^{(j)}.$$

On the other hand $\lambda_i(c_1\theta^{(i)} + c_j\theta^{(j)}) = c_i\lambda_i\theta^{(i)} + c_j\lambda_i\theta^{(j)} = 0$, so subtracting gives

$$c_j(\lambda_i - \lambda_j)\theta^{(j)} = 0.$$

Since $\theta^{(j)}$ is nonzero and $\lambda_i \neq \lambda_j$, we are forced to have $c_j = 0$, and similarly $c_i = 0$. This shows $\theta^{(i)}, \theta^{(j)}$ are orthogonal.

(2) $\theta^{(1)}$: for any θ we can express it as a linear combination $\sum_{i=1}^d c_i\theta^{(i)}$, so assuming each eigenvector is normalized,

$$\Sigma\theta = \sum_{i=1}^n c_i\lambda_i\theta^{(i)} \implies \|\Sigma\theta\| \leqslant \sum_{i=1}^d |c_i\lambda_i| \|\theta^{(i)}\| = \sum_{i=1}^d |c_i\lambda_i| \leqslant \sum_{i=1}^d |c_i|\lambda_1,$$

and this maximum is indeed obtained when $\theta = \theta^{(1)}$.

(3) Similar to (2) one can show that $\theta^{(d)}$ minimizes $\theta^T\Sigma\theta$. Note that $\theta_{\max} = \theta^{(1)}$ and $\theta_{\min} = \theta^{(d)}$, and by (1) we are done. $\square$