MATH 507a Homework 2

Qilin Ye

September 16, 2022

Problem 1: D1.6.5

Show that

(1) if $\epsilon > 0$, $\inf\{\mathbb{P}(|X| > \epsilon) : \mathbb{E}X = 0, \text{var}(X) = 1\} = 0$.

(2) if $y \geqslant 1$, $\sigma^2 \in (0, \infty)$, $\inf\{\mathbb{P}(|X| > y) : \mathbb{E}X = 1, \text{var}(X) = \sigma^2\} = 0$.

Proof. (1) Let $\delta > 0$ be given. We define a discrete random variable X by

$$\begin{cases} \mathbb{P}(X = 0) = 1 - \delta \\ \mathbb{P}(X = -1/\sqrt{\delta}) = \delta/2 \\ \mathbb{P}(X = 1/\sqrt{\delta}) = \delta/2. \end{cases} \tag{*}$$

Clearly $\mathbb{E}X = 0$ and $\mathbb{P}(|X| > \epsilon) \leqslant \delta/2 + \delta/2$. The variance is

$$\mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{1}{\delta} \cdot \frac{\delta}{2} + \frac{1}{\delta} \cdot \frac{\delta}{2} = 1.$$

Letting $\delta \downarrow 0$ we see $\mathbb{P}(|X| > \epsilon)$ can be made arbitrarily small so the infimum is 0.

(2) Call the random variable in (*) X_δ. Immediately we see $\sigma X_\delta + 1$ has expected value 1 and variance σ^2. By sending $\delta \to 0$ again, we see

$$\mathbb{P}(|X_\delta| > y) \leqslant \frac{\delta}{2} \to 0.$$ □

Problem 2: Durrett 1.6.10

Prove that

$$\mathbb{P}(\bigcup_{i=1}^n A_i) \leqslant \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \sum_{i<j<k} \mathbb{P}(A_i \cap A_j \cap A_k).$$

Proof. Let A be the union of A_n's and consider $f := 1_A$ as well as $g := \sum_{i=1}^n 1_{A_i} - \sum_{i<j} 1_{A_i \cap A_j} + \sum_{i<j<k} 1_{A_i \cap A_j \cap A_k}$. For $x \in \Omega$, let $k(x)$ be the number of A_i's that contain x. Then, if $k_x = 0$ then $f(x) = g(x) = 0$, and if $k_x > 0$,

$$f(x) = 1 \qquad \text{and} \qquad g(x) = \binom{k_x}{1} - \binom{k_x}{2} + \binom{k_x}{3}.$$

It is easy to check that if $k_x = 1$ then both sides are 1 and same when $k_x = 2$ or 3. When $k_x = 4$, $g(x) = 2 > f(x) = 1$, and when $k_x \geqslant 5$,

$$g(x) = k_x - \frac{k_x(k_x-1)}{2} + \frac{k_x(k_x-1)(k_x-2)}{6} = k_x + \frac{k_x(k_x-1)(k_x-5)}{6} \geqslant k_x > 1.$$

Therefore $g(x) \geqslant f(x)$, and taking expected value gives the claim. □

Problem 3: Durrett 1.6.13

Prove if $\mathbb{E}X_1^- < \infty$ and $X_n \uparrow X$, then $\mathbb{E}X_n \uparrow \mathbb{E}X$.

Proof. We decompose each X_n into $X_n^+ - X_n^-$. By assumption X_n^+ is increasing so by MCT, $\mathbb{E}X_n^+ \uparrow \mathbb{E}X^+$. On the other hand, the X_n^- are bounded by X_1^-, so by DCT, $\mathbb{E}X_n^- \downarrow \mathbb{E}X^-$. Hence $\mathbb{E}X_n \uparrow \mathbb{E}X$. □

Problem: A

Consider n independent trials of an experiment in which outcome a_i occurs with probability p_i for $i \leqslant m$. Let N be the number of outcomes a_i which do not occur at all during the n trials. Find $\mathbb{E}N$.

Solution. For $i \leqslant m$, let N_i be the indicator variable of whether a_i occurred at all. It follows from independence that $\mathbb{P}(N_i = 1) = (1-p_i)^n$. Therefore

$$\mathbb{E}N = \mathbb{E}\Big(\sum_{i=1}^m N_i\Big) = \sum_{i=1}^m (1-p_i)^n.$$

Problem: B

Suppose $\mathbb{E}|X| < \infty$ and $\mathbb{E}(X1_A) = 0$ for all $A \in \sigma(X)$. Prove $X = 0$ a.s.

Proof. Suppose not, i.e., $\mathbb{P}(X \neq 0) > 0$. Since $\mathbb{P}(X \neq 0) = \mathbb{P}(X > 0) + \mathbb{P}(X < 0)$ we WLOG assume the first one is positive. Note that

$$\{X > 0\} = \bigcup_{n\geqslant 1}\{X > 1/n\}$$

so there exists some n such that $\mathbb{P}(X > 1/n) > 0$. Note that $\{X > 1/n\} =: A$ is the preimage of $(-1/n, 1]$ which is in $\sigma(X)$, so by assumption $\mathbb{E}(X1_A) = 0$, but we have just shown that $\mathbb{E}(X1_A) > \mathbb{P}(X > 1/n)/n > 0$, contradiction. □

Problem: C

Suppose $Y \geqslant 0$ and $\mathbb{E}Y^2 < \infty$. Let $0 \leqslant a \leqslant \mathbb{E}Y$.

- Show that $\mathbb{E}(Y-a) \leqslant \mathbb{E}(Y1_{\{Y>a\}})$.
- Show that

$$\mathbb{P}(Y > a) \geqslant \frac{(\mathbb{E}Y - a)^2}{\mathbb{E}Y^2}.$$

Proof. • If $Y(\omega) \leqslant a$ then $Y(\omega) - a \leqslant 0 \leqslant Y(\omega)1_{\{Y>a\}}$. Similarly, if $Y(\omega) > a$ then $Y(\omega) - a \leqslant Y(\omega) = Y(\omega)1_{\{Y>a\}}$. Taking expected value gives the claim.

• Using the previous part and Cauchy Schwarz,

$$(\mathbb{E}Y - a)^2 \leqslant [\mathbb{E}(Y1_{\{Y>a\}})]^2 \leqslant \mathbb{E}Y^2\mathbb{E}(1^2_{\{Y>a\}}) = \mathbb{E}Y^2\mathbb{P}(Y > a).$$

□

Problem: D

Let $\{A_n, n \geqslant 1\}$ be events. Express the indicators of events $\limsup A_n$ and $\liminf A_n$ each in terms of the indicators of 1_{A_n}.

Solution. Since $x \in \limsup A_n$ if and only if there exists a increasing sequence of indices of which each set contains x and therefore $1_{A_n}(x) = 1$ for these sets, we have

$$1_{\limsup A_n} = \limsup 1_{A_n} \qquad \text{and similarly} \qquad 1_{\liminf A_n} = \liminf 1_{A_n}.$$

Problem: E

Suppose $X \geqslant 0$, $p > 0$, and $\mathbb{E}X^p < \infty$. Show that $t^p\mathbb{P}(X > t) \to 0$ as $t \to \infty$.

Proof. Note that

$$t^p\mathbb{P}(X > t) = t^p\mathbb{P}(X^p > t^p) = \mathbb{E}[t^p 1_{\{X^p>t^p\}}].$$

For almsot every ω, $1_{X^p>t^p}(\omega) = 0$ for sufficiently large t, so $\lim_{t\to\infty} t^p 1_{\{X^p>t^p\}} \equiv 0$ almost surely. Also that $t^p 1_{\{X^p>t^p\}} \leqslant X^p$ for all t: if $X(\omega) \leqslant t$ then $X^p(\omega) \leqslant t^p$ and the LHS is $0 \leqslant X^p(\omega)$; if $X(\omega) > t$ then the LHS is $t^p < X^p(\omega)$. Therefore, by DCT, $\mathbb{E}[t^p 1_{\{X^p>t^p\}}] = t^p\mathbb{P}(X > t) \to \mathbb{E}[0] = 0$. □

Problem: F

- Show that a convex function $\varphi : \mathbb{R} \to \mathbb{R}$ is strictly convex iff there is no interval $[a, b]$ of positive length on which φ is linear.
- Suppose φ is strictly convex and X is a r.v. with $\mathbb{E}|X| < \infty$, $\mathbb{E}|\varphi(X)| < \infty$, and $\mathbb{E}\varphi(X) = \varphi(\mathbb{E}X)$. Show that $X = \mathbb{E}X$ a.s.

Proof. • The easy direction first: suppose φ is linear on $[a, b]$. Then $\varphi((a + b)/2) = (\varphi(a) + \varphi(b))/2$, so φ is not stictly convex.

Conversely, suppose φ is not strictly convex. By definition that means there exist $a < b$ and $\lambda \in (0, 1)$ with $\varphi(\lambda a + (1 - \lambda)b) = \lambda\varphi(a) + (1 - \lambda)\varphi(b)$. For convenience denote $\lambda a + (1 - \lambda)b$ as c. By convexity, there exists a linear function ℓ with $\ell \leqslant \varphi$ and $\ell(c) = \varphi(c)$. Since $(a, \varphi(a)), (c, \varphi(c))$, and $(b, \varphi(b))$ are colinear, either $\varphi(a) - \ell(a)$ and $\varphi(b) - \ell(b)$ are both 0, or (exactly) one between them is negative. The latter cannot happen by ℓ's construction. Since convex functions from $\mathbb{R}$ to $\mathbb{R}$ are continuous, so is $\varphi - \ell$. Therefore, by IVT, either $\varphi - \ell \equiv 0$ on $[a, b]$ (in which case we prove the claim) or there exists $d \in [a, b]$ with $\varphi(d) < \ell(d)$, contradiction, also completing the proof.

- Let ℓ be a global underestimator of φ with $\ell(c) = \varphi(c)$ where $c = \mathbb{E}X$. Let its equation be $\ell(X) = aX + b$. It follows that
$$\int \varphi(X)\,\mathrm{d}\mathbb{P} \geqslant \int aX + b\,\mathrm{d}\mathbb{P} = a\mathbb{E}X + b = \ell(\mathbb{E}X) = \varphi(\mathbb{E}X).$$
The equality holds if and only if $\varphi(X) = aX + b$ almost surely, that is, φ is affine a.s. This contradicts the previous part unless we trivialise the problem by requiring $X = \mathbb{E}X$ a.s. □

- Suppose X is not equal to $\mathbb{E}X$ a.s. Let ℓ be a global underestimator of φ with $\ell(\mathbb{E}X) = \varphi(\mathbb{E}X)$. Let its equation be $\ell(X) = aX + b$. It follows from strict convexity that
$$\varphi(x) - \varphi(\mathbb{E}X) > a(x - \mathbb{E}X)$$
for all $x \neq \mathbb{E}X$. If $X \neq \mathbb{E}X$ a.s. then integrating the above inequality gives
$$\int_\Omega \varphi(x) - \varphi(\mathbb{E}X)\,\mathrm{d}\mathbb{P} = \int_{\{X \neq \mathbb{E}X\}} \varphi(x) - \varphi(\mathbb{E}X)\,\mathrm{d}\mathbb{P} = \mathbb{E}\varphi(X) - \varphi(\mathbb{E}X) > a(\mathbb{E}X - \mathbb{E}X) = 0,$$
so $\mathbb{E}\varphi(X) > \varphi(\mathbb{E}X)$. This proves the claim.