521 Homework 10

Daniel Rui - 12/4/19

Problem 1

We have independent $X_1, X_2, \dots$ with $P([X_n = 1]) = p_n$ and $P([X_n = 0]) = 1 - p_n$.

(i) We want to show that $X_n \to_p 0 \iff p_n \to 0$.

($\implies$) From the definition of $\to_p$, we know that $\forall \epsilon_1, \epsilon_2, \exists N$ s.t. $n \geq N \implies P([|X_n| > \epsilon_1]) < \epsilon_2$. So then $n \geq N \implies p_n = P([X_n = 1]) \leq P([|X_n| > \epsilon_1]) < \epsilon_2$, or in other words, $\forall \epsilon_2 > 0, \exists N$ s.t. $n \geq N \implies p_n < \epsilon_2$, which is exactly the definition of $p_n \to 0$.

($\impliedby$) Again from definition, we know that $\forall \epsilon, \exists N$ s.t. $n \geq N \implies p_n < \epsilon \iff -\epsilon < -p_n$ and so $1 - \epsilon < 1 - p_n = P([X_n = 0]) \leq P([|X_n| \leq \epsilon'])$, for all $\epsilon' > 0$, which means that $1 - P([|X_n| \leq \epsilon']) < \epsilon \iff P([|X_n| > \epsilon']) < \epsilon$ and so by definition $X_n \to_p 0$.

(ii) We want to show that $X_n \to_{\text{a.s.}} 0 \iff \sum_{n=1}^{\infty} p_n < \infty$.

($\implies$) Assume not. Then, $\sum_{n=1}^{\infty} p_n = \sum_{n=1}^{\infty} P([X_n = 1]) = \infty$, and so Borel-Cantelli tells us that $P(A := \{\omega \in \Omega : X_n(\omega) = 1 \text{ for infinitely many } n\}) = 1$. But from the definition of $\to_{\text{a.s.}}$, we know that $P(B := \{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = 0\}) = 1$ as well. But A and B are disjoint sets, because if $\omega \in A$, then $X_n(\omega)$ could not possibly converge to 0 because for every N, there's $n' \geq N$ s.t. $X_{n'}(\omega) = 1$! Similarly, $\omega \in B \implies \omega \notin A$. Thus, $P(A \cup B) = P(A) + P(B) = 2$, which is clearly impossible.

($\impliedby$) $\sum_{n=1}^{\infty} p_n = \sum_{n=1}^{\infty} P([X_n = 1]) < \infty$, and so by Borel-Cantelli, we have that

$$\begin{aligned}
&P(\{\omega \in \Omega : X_n(\omega) = 1 \text{ for infinitely many } n\}) = 0 \\
\implies &P(\{\omega \in \Omega : X_n(\omega) = 1 \text{ for finitely many } n\}) = 1 \\
\implies &P(\{\omega \in \Omega : \exists N_\omega \text{ s.t. } n \geq N_\omega \implies X_n(\omega) = 0\}) = 1 \\
\implies &P(\{\omega \in \Omega : X_n(\omega) \to 0\}) = 1 \\
\implies &X_n \to_{\text{a.s.}} 0
\end{aligned}$$

Problem 2

We have $X_1, X_2, \dots$ i.i.d. Exponential(1) so fixing an $\epsilon > 0$, we have that $P([X_n > (\log n)(1+\epsilon)]) = e^{-(\log n)(1+\epsilon)} = n^{-(1+\epsilon)}$ and so

$$\sum_{n=1}^{\infty} P([X_n > (\log n)(1+\epsilon)]) = \sum_{n=1}^{\infty} \frac{1}{n^{1+\epsilon}}$$

which we know converges (by say the integral test). Thus, Borel-Cantelli says that

$$\begin{aligned}
&P(\{\omega \in \Omega : X_n(\omega) > (\log n)(1+\epsilon) \text{ for infinitely many } n\}) = 0 \\
\implies &P(\{\omega \in \Omega : X_n(\omega) > (\log n)(1+\epsilon) \text{ for finitely many } n\}) = 1 \\
\implies &P(\{\omega \in \Omega : \exists N_\omega \text{ s.t. } n \geq N_\omega \implies X_n(\omega) \leq (\log n)(1+\epsilon)\}) = 1 \\
\implies &P\left(\left\{\omega \in \Omega : \exists N_\omega \text{ s.t. } n \geq N_\omega \implies \frac{X_n(\omega)}{\log n} \leq (1+\epsilon)\right\}\right) = 1 \\
\implies &\limsup_{n\to\infty} \frac{X_n}{\log n} \leq 1+\epsilon \quad \text{a.s.} \quad (\forall \epsilon > 0) \\
\implies &\limsup_{n\to\infty} \frac{X_n}{\log n} \leq 1 \quad \text{a.s.}
\end{aligned}$$

Similarly, fixing an $\epsilon > 0$, we have that $P([X_n > (\log n)(1-\epsilon)]) = e^{-(\log n)(1-\epsilon)} = n^{-(1-\epsilon)}$ and so

$$\sum_{n=1}^{\infty} P([X_n > (\log n)(1-\epsilon)]) = \sum_{n=1}^{\infty} \frac{1}{n^{1-\epsilon}}$$

which we know diverges (by say the integral test). Thus, Borel-Cantelli says that

$$\begin{aligned}
&P(\{\omega \in \Omega : X_n(\omega) > (\log n)(1-\epsilon) \text{ for infinitely many } n\}) = 1 \\
\implies &P(\{\omega \in \Omega : \forall N, \exists n \geq N \text{ s.t. } X_n(\omega) > (\log n)(1-\epsilon)\}) = 1 \\
\implies &P\left(\left\{\omega \in \Omega : \forall N, \exists n \geq N \text{ s.t. } \frac{X_n}{\log n} > (1-\epsilon)\right\}\right) = 1 \\
\implies &\limsup_{n\to\infty} \frac{X_n}{\log n} > 1-\epsilon \quad \text{a.s.} \quad (\forall \epsilon > 0) \\
\implies &\limsup_{n\to\infty} \frac{X_n}{\log n} \geq 1 \quad \text{a.s.}
\end{aligned}$$

and so the two directions tell us that $\limsup_{n\to\infty} \frac{X_n}{\log n} = 1$.

Problem 4

We are given $\xi_1, \xi_2, \ldots$ and $\Theta_1, \Theta_2, \ldots$ are all independent and i.i.d. Uniform(0,1), and a continuous function $h : [0,1] \to [0,1]$.

(a) Then, $X_n := 1_{[h(\xi_n) > \Theta_n]}$ are independent and

$$\begin{aligned}
\mathbb{E}[X_n] = P([h(\xi_n) > \Theta_n]) &= \int_0^1 P([h(\xi_n) > \Theta_n] | \Theta_n = x) f_{\Theta_n}(x)\, dx \\
&= \int_0^1 P([h(\xi_n) > x])\, dx = \int_0^\infty P([h(\xi_n) > x])\, dx = \mathbb{E}[h(\xi_n)] \\
&= \int_0^1 h(x)\, dF_{\xi_n}(x) = \int_0^1 h(x)\, d\lambda(x) = \int_0^1 h(x)\, dx
\end{aligned}$$

where the second equality follows from the law of total probability (or expectation). I don't think we covered it in class, so I don't know the proof, but I suppose I will just use it here. The sixth equality follows from the law of the unconscious statistician. Thus, by the strong law of large numbers,

$$\bar{X}_n \to_{\text{a.s.}} \int_0^1 h(x)\,dx$$

(b) Defining $Y_n := h(\xi_n)$, we see that $\mathbb{E}[Y_n] = \mathbb{E}[\xi_n]$ which we saw from above to be $\int_0^1 h(x)\,dx$. Thus again by SLLN,

$$\bar{Y}_n \to_{\text{a.s.}} \int_0^1 h(x)\,dx$$

(c) First we evaluate $\operatorname{Var}[X_n]$ and $\operatorname{Var}[Y_n]$:

$$\operatorname{Var}[X_n] = \mathbb{E}\left[X_n^2\right] - (\mathbb{E}[X_n])^2 = \mathbb{E}[X_n] - (\mathbb{E}[X_n])^2 := \mu - \mu^2$$

(because X_n is only 0 or 1, $X_n = X_n^2$) and so

$$\operatorname{Var}\left[\bar{X}_n\right] = \frac{1}{n^2}\sum_{i=1}^n \operatorname{Var}[X_i] = \frac{1}{n}\operatorname{Var}[X_n] = \frac{1}{n}(\mu - \mu^2)$$

Similarly

$$\operatorname{Var}[Y_n] = \mathbb{E}\left[Y_n^2\right] - (\mathbb{E}[Y_n])^2 = \mathbb{E}\left[h^2(\xi_n)\right] - (\mathbb{E}[Y_n])^2 = \mathbb{E}\left[h^2(\xi_n)\right] - \mu^2$$

and so

$$\operatorname{Var}\left[\bar{Y}_n\right] = \frac{1}{n^2}\sum_{i=1}^n \operatorname{Var}[Y_i] = \frac{1}{n}\operatorname{Var}[Y_n] = \frac{1}{n}(\mathbb{E}\left[h^2(\xi_n)\right] - \mu^2) = \frac{1}{n}\int_0^1 h^2(x)\,dx - \frac{\mu^2}{n}$$

But because h only goes to $[0,1]$, $h^2(x) \leq h(x)$, and so $\int_0^1 h^2(x)\,dx \leq \mu$ and so $\operatorname{Var}\left[\bar{Y}_n\right] \leq \operatorname{Var}\left[\bar{X}_n\right]$.

521 Homework 9

Daniel Rui - 11/27/19

Problem 1

(a) We want to show that A and B are independent events $\iff$ $\sigma[A]$ and $\sigma[B]$ are independent sigma algebras, i.e. $P(A \cap B) = P(A) \cdot P(B) \iff P(A' \cap B') = P(A') \cdot P(B')$ for all $A' \in \sigma[A] = \{\varnothing, A, A^\complement, \Omega\}$ and $B' \in \sigma[B] = \{\varnothing, B, B^\complement, \Omega\}$.

($\implies$) If one of the sets in question is $\varnothing$, then for any set S, $P(\varnothing \cap S) = 0 = P(\varnothing) \cdot P(S)$; similarly if one of the sets in question is Ω, then for any set S, $P(\Omega \cap S) = P(S) = P(\Omega) \cdot P(S)$. This leaves three cases:

- Reminder: $P(A \cap B) + P(A \cap B^\complement)$. Thus, $P(A \cap B^\complement) = P(A) - P(A \cap B) = P(A) - P(A) \cdot P(B) = P(A)\big(1 - P(B)\big) = P(A) \cdot P(B^\complement)$. The case $P(A^\complement \cap B) = P(A^\complement) \cdot P(B)$ follows by symmetry.
- $P(A^\complement \cap B^\complement) = P(A^\complement) - P(A^\complement \cap B) = P(A^\complement) - P(A^\complement) \cdot P(B) = P(A^\complement)\big(1 - P(B)\big) = P(A^\complement) \cdot P(B^\complement)$, where the second equality follows from the above bullet point.

($\impliedby$) In particular, $A \in \sigma[A]$ and $B \in \sigma[B]$, so obviously $P(A \cap B) = P(A) \cdot P(B)$.

(b) We again want to equate two definitions of independence: for every $k \in \{2, \ldots, n\}$, P(the intersection of *any* k of $\{A_1, \ldots, A_n\}$) = the product of the corresponding $\{P(A_1), \ldots, P(A_n)\}$ $\iff$ $\sigma[A_1], \ldots, \sigma[A_n]$ are independent sigma fields, i.e. $P(A_1' \cap \ldots \cap A_n') = \prod_{i=1}^n P(A_i')$ for all $A_i' \in \sigma[A_i]$.

($\implies$) Like above, the cases when $A_i' = \varnothing$ are trivial, and the cases where $A_i' = \Omega$ just simplify to proving the statement for any k sets $A_i' = A_i, A_i^\complement$ for all $k \in \{2, \ldots, n\}$, i.e. proving that for every $k \in \{2, \ldots, n\}$, P(the intersection of any k of $\{A_1' := A_1$ or $A_1^\complement, \ldots, A_n' := A_n'$ or $A_n^\complement\}$) = the product of the corresponding $\{P(A_1'), \ldots, P(A_n')\}$.

we proceed by induction: we can apply the argument from (a) that for any two (distinct) indices $1 \le i_1, i_2 \le n$, $P(A_{i_1}' \cap A_{i_2}') = P(A_{i_1}')P(A_{i_2}')$. Now we assume the statement holds for any $k-1$ sets. The key observation is that *if* we have k sets $A_1', \ldots, A_k'$ that satisfy $P(A_1' \cap \ldots \cap A_n') = P(A_1') \cdots P(A_n')$, then replacing any A_i' with $A_i'^\complement$ preserves the equality:

$$\begin{aligned}
P(A_1' \cap \ldots \cap A_i'^\complement \cap \ldots \cap A_k') &= P(A_1' \cap \ldots \cap \Omega \cap \ldots \cap A_k') - P(A_1' \cap \ldots \cap A_i' \cap \ldots \cap A_k') \\
&= P(\underbrace{A_1' \cap \ldots \cap A_{i-1}' \cap A_{i+1}' \cap \ldots \cap A_k'}_{k-1 \text{ sets}}) - P(A_1') \cdots P(A_k') \\
&= P(A_1') \cdots P(A_{i-1}') \cdot P(A_{i+1}') \cdots P(A_k') - P(A_1') \cdots P(A_k') \\
&= \big(P(A_1') \cdots P(A_{i-1}') \cdot P(A_{i+1}') \cdots P(A_k')\big) \cdot \big(1 - P(A_i')\big) \\
&= P(A_1') \cdots P(A_{i-1}') \cdot P(A_i'^\complement) \cdot P(A_{i+1}') \cdots P(A_k')
\end{aligned}$$

Thus given any configuration of k sets with k' of them being $A_i' = A_i$ and $k - k'$ of them being $A_i' = A_i^\complement$, we can always start from our given: $P(A_{i_1} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k})$ and toggle each of the $k - k'$ sets to its complement to get the desired equality.

($\Longleftarrow$) Given any set I of k indices from $[n] := \{1, \ldots, n\}$, denoting $I^{\complement} := [n] \setminus I$, then we know from the independence of the sigma fields that

$$P\left(\bigcap_{i \in I} A_i\right) = P\left(\bigcap_{i \in I} A_i \cap \bigcap_{i \in I^{\complement}} \Omega\right) = \prod_{i \in I} P(A_i) \cdot \prod_{i \in I^{\complement}} P(\Omega) = \prod_{i \in I} P(A_i) \cdot 1^{n-k} = \prod_{i \in I} P(A_i)$$

Problem 2

Take $\Omega = \{1, 2, 3, 4\}$, $\mathcal{A}_1 = \big\{\{1, 2\}, \{1, 3\}\big\}$, and $\mathcal{A}_2 = \big\{\{2, 3\}\big\}$ (and of course $P = \#/4$). Then, $P(\{1, 2\} \cap \{2, 3\}) = P(\{3\}) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(\{1, 2\}) \cdot P(\{2, 3\})$, so $\{1, 2\}$ and $\{2, 3\}$ are independent events. A similar process yields that $\{1, 3\}$ and $\{2, 3\}$ are independent; thus, $\mathcal{A}_1$ and $\mathcal{A}_2$ are independent collections.

However, $\{2\} = \left(\{1, 3\} \cup \{1, 2\}^{\complement}\right)^{\complement}$ is in $\sigma[\mathcal{A}_1]$, and $P(\{2\} \cap \{2, 3\}) = \frac{1}{4} \neq \frac{1}{4} \cdot \frac{1}{2} = P(\{2\}) \cdot P(\{2, 3\})$, so $\sigma[\mathcal{A}_1]$ and $\sigma[\mathcal{A}_2]$ are not independent.

Problem 3

We have a sequence of random variables X_n (we will assume that past a certain point C, all the X_n are only infinite on a set N_n of measure 0 — otherwise e.g. you could just take $X_1, X_2, \ldots = \infty$ and there would exist no such c_n). We prove a quick lemma first: given $X \geq 0$ s.t. X is infinite only on a set N of measure 0, for every $\epsilon > 0$, we can find K s.t. $P([X > K]) < \epsilon$. Define $A_n = [X > n]$; then the A_n form a monotone decreasing sequence. Thus, we can use the limit-measure interchange theorems to see that

$$\lim_{n \to \infty} P(A_n) = P\left(\lim_{n \to \infty} A_n\right) = P\left(\bigcap_{n=1}^{\infty} A_n\right) = P(N) = 0$$

which concludes the proof.

Now for any X_n (for $n > C$, the "certain point" mentioned above), choose $\epsilon_n = 2^{-n}$, and find the associated K_n. Define $Y_n = \dfrac{X_n}{nK_n}$. Then,

$$P\left(\left[|Y_n| > \frac{1}{n}\right]\right) = P\left(\left[\left|\frac{X_n}{nK_n}\right| > \frac{1}{n}\right]\right) = P([|X_n| > K_n]) < \epsilon_n = \frac{1}{2^n}$$

With this construction, we now see that

$$\sum_{n=1}^{\infty} P\left(\left[|Y_n| > \frac{1}{n}\right]\right) < \sum_{n=1}^{C} P(\Omega) + \sum_{n=C+1}^{\infty} \frac{1}{2^n} = C + \frac{1}{2^C} < \infty$$

and so by Borel-Cantelli, we know that $P\left(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} [|Y_n| > \frac{1}{n}]\right) = 0$, which leads to the following implications:

$$\implies P(\{\omega \in \Omega : |Y_n(\omega)| > \frac{1}{n} \text{ for infinitely many } n\}) = 0$$
$$\implies P(\{\omega \in \Omega : |Y_n(\omega)| > \frac{1}{n} \text{ for finitely many } n\}) = 1$$
$$\implies P(\{\omega \in \Omega : \exists N_\omega \text{ s.t. } n \geq N_\omega \implies |Y_n(\omega)| \leq \frac{1}{n}\}) = 1$$
$$\implies P(\{\omega \in \Omega : Y_n(\omega) \to 0\}) = 1$$
$$\implies Y_n \to_{\text{a.s.}} 0$$

Thus, the c_n s.t. $\dfrac{X_n}{c_n} \to_{\text{a.s.}} 0$ is simply the nK_n from above.

Problem 5

We have independent $X_1, X_2, \ldots$, and we want to prove that $\sup_{n\in\mathbb{N}} X_n(\omega) < \infty$ for almost every $\omega \in \Omega \iff \sum_{n=1}^{\infty} P([X_n > M]) < \infty$ for some $M < \infty$ (assuming that *all* the X_n are only infinite on a set A_n of measure 0 — otherwise e.g., if $X_1 = \infty$ and $X_2, X_3, \ldots = 0$, the result is clearly false).
($\implies$) Define $B_k = \{\omega \in \Omega : \sup_{n\in\mathbb{N}} X_n(\omega) \leq k\}$; then the B_k form a monotone increasing sequence, and so we can apply the limit-measure interchange theorems:

$$\lim_{k\to\infty} P(B_k) = P\left(\left\{\omega \in \Omega : \sup_{n\in\mathbb{N}} X_n(\omega) < \infty\right\}\right) = 1$$

so in other words, fixing an $\epsilon > 0$, there exists some K s.t.

$$k \geq K \implies P(B_k) > 1 - \epsilon \implies P(B_k^{\complement}) < \epsilon \implies P\left(\left[\sup_{n\in\mathbb{N}} X_n > k\right]\right) < \epsilon$$

and in particular for $k = K$,

$$P(\{\omega \in \Omega : \infty\text{-many } X_n(\omega) > K\}) \leq P\left(\left[\sup_{n\in\mathbb{N}} X_n > K\right]\right) < \epsilon$$

Now assume for sake of contradiction that $\sum_{n=1}^{\infty} P([X_n > K])$ is actually infinite; then Borel-Cantelli would tell us that $P([X_n > K \text{ i.o.}]) = 1$, which is incompatible with our finding above that it's less than the fixed ϵ; therefore, $\sum_{n=1}^{\infty} P([X_n > K]) < \infty$.
($\impliedby$) Stealing one of the implications from the end of Problem 3, Borel-Cantelli gives that

$$\sum_{n=1}^{\infty} P([X_n > M]) < \infty \implies P(A := \{\omega \in \Omega : \exists N_\omega \text{ s.t. } n \geq N_\omega \implies X_n(\omega) \leq M\}) = 1$$

Notice that $P(A^{\complement} \cup A_1 \cup A_2 \cup \ldots) < P(A^{\complement}) + P(A_1) + P(A_2) + \ldots = 0$, and so $P\big((A^{\complement} \cup A_1 \cup A_2 \cup \ldots)^{\complement}\big) = P(A \cap A_1^{\complement} \cap A_2^{\complement} \cap \ldots) = 1$. Thus, for every $\omega \in (A \cap A_1^{\complement} \cap A_2^{\complement} \cap \ldots)$, i.e. almost surely, $\sup_{n\in\mathbb{N}} X_n(\omega)$ is simply $\max\{X_1(\omega), \ldots, X_{N_\omega - 1}(\omega), M\}$, where we can take the maximum because we only have a finite number of values (N_ω many, to be exact) to deal with. Because these values themselves are all finite, we have that $\sup_{n\in\mathbb{N}} X_n(\omega)$ is finite (almost surely).

521 Homework 8

Daniel Rui - 11/20/19

Problem 1

(a) We know that $P(X > x) = \int_\Omega 1_{[X>x]}\, dP$ and that for any $f \geq 0$, $\int_\mathbb{R} f(\omega)\, d\lambda = \int_{-\infty}^{\infty} f(x)\, dx$, so

$$\int_0^\infty P(X > x)\, dx = \int_\mathbb{R} 1_{[x>0]} P(X > x)\, d\lambda(x) = \int_\mathbb{R} 1_{[x>0]} \left[\int_\Omega 1_{[X>x]}\, dP\right] d\lambda(x)$$

And by Tonelli (which we can apply because indicator functions are ≥ 0), we can interchange integrals. Up to this point, we've been defining our indicator functions in a manner that seems to suggest that we are fixing x and considering sets of $\omega \in \Omega$, so we will now explicitly write out the two-variable function inside the double integral:

$$Y(\omega, x) = \begin{cases} 1 & \text{if } X(\omega) > x > 0 \\ 0 & \text{otherwise} \end{cases}$$

and so the computation is as follows:

$$\begin{aligned} \int_0^\infty P(X > x)\, dx &= \int_\mathbb{R}\int_\Omega Y(\omega, x)\, dP(\omega)\, d\lambda(x) = \int_\Omega \int_\mathbb{R} Y(\omega, x)\, d\lambda(x)\, dP(\omega) \\ &= \int_\Omega \int_0^{X(\omega)} 1\, d\lambda(x)\, dP(\omega) = \int_\Omega X(\omega)\, dP(\omega) = \mathbb{E}[X] \end{aligned}$$

Finally, because $P(X > x) = 1 - F(x)$, we can replace $P(X > x)$ with $1 - F(x)$ in the above integrals.

(b) Because we have that $\mathbb{E}[|X|] < \infty$, $\mathbb{E}[X^+] < \infty$ and $\mathbb{E}[X^-] < \infty$, so we can safely say that $\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-]$. From part (a), we know that

$$\mathbb{E}\left[X^+\right] = \int_0^\infty P(X^+ > x)\, dx = \int_0^\infty P(X > x)\, dx$$

and similarly

$$\mathbb{E}\left[X^-\right] = \int_0^\infty P(X^- > x)\, dx = \int_0^\infty P(X < -x)\, dx = \int_{-\infty}^0 P(X < x)\, dx$$

Thus, taking the difference yields

$$\mathbb{E}[X] = \int_0^\infty P(X > x)\, dx - \int_{-\infty}^0 P(X < x)\, dx = \int_0^\infty (1 - F(x))\, dx - \int_{-\infty}^0 F(x)\, dx$$

Problem 3

We know that $F(x+t) = P(X \le x+t)$ and $F(x) = P(X \le x)$. Hence,

$$F(x+t) - F(x) = P(x < X \le x+t) = \int_\Omega 1_{[x<X\le x+t]}\, dP$$

Similar to Problem 1, define

$$Y(\omega, x) = \begin{cases} 1 & \text{if } x < X(\omega) \le x+t \\ 0 & \text{otherwise} \end{cases}$$

and because $x < X(\omega) \le x+t \iff X(\omega) - t < x < X(\omega)$, have that

$$\begin{aligned} \int_\mathbb{R} P(x < X \le x+t)\, dx &= \int_\mathbb{R} \int_\Omega Y(\omega, x)\, dP(\omega)\, d\lambda(x) = \int_\Omega \int_\mathbb{R} Y(\omega, x)\, d\lambda(x)\, dP(\omega) \\ &= \int_\Omega \int_{X(\omega)-t}^{X(\omega)} 1\, d\lambda(x)\, dP(\omega) = \int_\Omega X(\omega) - (X(\omega) - t)\, dP(\omega) \\ &= \int_\Omega t\, dP(\omega) = t \cdot P(\Omega) = t \end{aligned}$$

521 Homework 7

Daniel Rui - 11/13/19

Problem 1

We have the "prototypical" signed measure $\phi(A) = \int_A X \, d\mu$ (with the restriction that $X^- \in \mathcal{L}^1$).

Define Ω^+ to be $\{\omega \in \Omega : X(\omega) \geq 0\}$, and similarly Ω^- to be $\{\omega \in \Omega : X(\omega) < 0\}$. Clearly, $\Omega^+ \cup \Omega^- = \Omega$ and $\Omega^+ \cap \Omega^- = \varnothing$, and for any $A \in \Omega^+$, $\int_A X \, d\mu$ will obviously be ≥ 0. Likewise for $A \in \Omega^-$, $\int_A X \, d\mu$ will be ≤ 0. From this, we can say that

$$\phi^+(A) = \phi(A \cap \Omega^+) = \int_{A \cap \Omega^+} X \, d\mu = \int_A X^+ \, d\mu \quad \text{and similarly} \quad \phi^-(A) = \int_A X^- \, d\mu$$

and so

$$|\phi|(A) = \phi^+(A) + \phi^-(A) = \int_A X^+ + X^- \, d\mu = \int_A |X| \, d\mu \implies |\phi|(\Omega) = \int_\Omega |X| \, d\mu = \mathbb{E}\left[|X|\right]$$

Problem 2

We are given a σ-finite measure μ and a finite measure ν. We then define a function $\phi : \mathcal{A} \to (-\infty, \infty]$ as $\phi(A) = \mu(A) - \nu(A)$. Note in particular that we defined ν to be finite in order to exclude $-\infty$.

(a) ϕ is a signed measure: $\phi(\varnothing) = \mu(\varnothing) - \nu(\varnothing) = 0 - 0 = 0$; and for disjoint sets, $\{A_n\}$,

$$\phi\left(\bigcup_{i=1}^{\infty} A_i\right) = \mu\left(\bigcup_{i=1}^{\infty} A_i\right) - \nu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i) - \sum_{i=1}^{\infty} \nu(A_i) = \sum_{i=1}^{\infty} (\mu(A_i) - \nu(A_i)) = \sum_{i=1}^{\infty} \phi(A_i)$$

where we can pull in the ν sum because it is convergent. All the more reason to throw out $-\infty$!

(b) If we have that $[\mu + \nu](A) = 0$, then $\mu(A)$ and $\nu(A)$ must both equal 0. This means that $\mu \ll \mu + \nu$ and $\nu \ll \mu + \nu$. Furthermore, having both $\mu(A)$ and $\nu(A)$ be 0 would imply that $\phi(A) = \mu(A) - \nu(A) = 0$, so $\phi \ll \mu + \nu$. By the Radon-Nikodym theorem, we have that there exists f, g, where $g \in \mathcal{L}^1$ w.r.t. $[\mu + \nu]$, such that

$$\mu(A) = \int_A f \, d[\mu + \nu] \quad \text{and} \quad \nu(A) = \int_A g \, d[\mu + \nu] \implies \phi(A) = \int_A (f - g) \, d[\mu + \nu],$$

but also that there is also some measurable X such that $\phi(A) = \int_A X \, d[\mu + \nu]$. Radon-Nikodym assures us that $X = f - g$ a.e. w.r.t. $[\mu + \nu]$.

(c) From Problem 1, we have that $\phi^+(A) = \int_A (f - g)^+ \, d[\mu + \nu]$, $\phi^-(A) = \int_A (f - g)^- \, d[\mu + \nu]$, $|\phi|(A) = \int_A |f - g| \, d[\mu + \nu]$, and in the case that μ is finite, $f \in \mathcal{L}^1[\mu + \nu]$, so $|\phi|(\Omega) = \mathbb{E}\left[|f - g|\right] < \infty$

Problem 3

We define the total variation distance between two probability measures P, Q as

$$d_{TV}(P,Q) = \sup_{A\in\mathcal{A}} |P(A) - Q(A)|$$

Now if we have any measure μ that satisfies $P \ll \mu$ and $Q \ll \mu$, then (by Radon-Nikodym) there is some measurable p, q s.t.

$$P(A) = \int_A p\, d\mu \quad \text{and} \quad Q(A) = \int_A q\, d\mu$$

(a) We want to prove that the following identity holds (for any μ dominating P, Q)

$$d_{TV}(P,Q) = \frac{1}{2}\int_\Omega |p-q|\, d\mu$$

Let a set $G = [p - q \geq 0]$ (G for "greater than") and a set $L = [p-q<0] = [q-p>0]$. Then,

$$\begin{aligned}
\int_\Omega |p-q|\, d\mu &= \int_G (p-q)\, d\mu + \int_L (q-p)\, d\mu \\
&\leq \sup_{A\in\mathcal{A}}\left\{\int_A (p-q)\, d\mu\right\} + \sup_{A\in\mathcal{A}}\left\{\int_A (q-p)\, d\mu\right\} \\
&\leq \sup_{A\in\mathcal{A}}\left\{\left|\int_A (p-q)\, d\mu\right|\right\} + \sup_{A\in\mathcal{A}}\left\{\left|-\int_A (p-q)\, d\mu\right|\right\} \\
&= 2\sup_{A\in\mathcal{A}}\left\{\left|\int_A (p-q)\, d\mu\right|\right\}
\end{aligned}$$

Now notice that

$$\int_\Omega (p-q)\, d\mu = \int_\Omega p\, d\mu - \int_\Omega q\, d\mu = P(\Omega) - Q(\Omega) = 1 - 1 = 0$$

which implies that

$$\int_G (p-q)\, d\mu + \int_L (p-q)\, d\mu = 0 \implies \int_G (p-q)\, d\mu = \int_L (q-p)\, d\mu$$

And thus for any $A \in \mathcal{A}$, we know that

$$\begin{aligned}
\left|\int_A (p-q)\, d\mu\right| &= \max\left\{\int_A (p-q)\, d\mu, -\int_A (p-q)\, d\mu\right\} \\
&= \max\left\{\int_A (p-q)\, d\mu, \int_A (q-p)\, d\mu\right\} \\
&\leq \max\left\{\int_{A\cap G} (p-q)\, d\mu, \int_{A\cap L} (q-p)\, d\mu\right\} \\
&\leq \max\left\{\int_G (p-q)\, d\mu, \int_L (q-p)\, d\mu\right\} \\
&= \int_G (p-q)\, d\mu = \int_L (q-p)\, d\mu = \frac{1}{2}\int_\Omega |p-q|\, d\mu
\end{aligned}$$

Because this holds for all $A \in \mathcal{A}$, it holds for the supremum as well, so

$$\frac{1}{2}\int_\Omega |p-q|\,d\mu \leq \sup_{A\in\mathcal{A}}\left\{\left|\int_A (p-q)\,d\mu\right|\right\} \leq \frac{1}{2}\int_\Omega |p-q|\,d\mu$$

which means they are equal. Again, the proof holds for any arbitrary measure μ that satisfies $P, Q \ll \mu$.

(b) Trivially from the above definitions of P, Q, we have that

$$[P-Q](A) = \int_A (p-q)\,d\mu$$

From part (b), we know that

$$d_{TV}(P,Q) = \frac{1}{2}\int_\Omega |p-q|\,d\mu = \frac{1}{2}\mathbb{E}\left[|p-q|\right] = \frac{1}{2}|[P-Q]|(\Omega)$$

where the third equality follows from Problem 2 part (c).

Problem 4

Let $A = [X \geq 0]$ and $B = [X < 0]$. Note that on A, $X = |X|$ and on B, $X = -|X|$; and on Ω^+, $\phi = |\phi|$, and on Ω^-, $\phi = -|\phi|$. Now observe:

$$\begin{aligned}
\left|\int_\Omega X\,d\phi\right| &= \left|\int_{A\cap\Omega^+} X\,d\phi + \int_{B\cap\Omega^+} X\,d\phi + \int_{A\cap\Omega^-} X\,d\phi + \int_{B\cap\Omega^-} X\,d\phi\right| \\
&= \left|\int_{A\cap\Omega^+} |X|\,d\phi + \int_{B\cap\Omega^+} -|X|\,d\phi + \int_{A\cap\Omega^-} |X|\,d\phi + \int_{B\cap\Omega^-} -|X|\,d\phi\right| \\
&= \left|\int_{A\cap\Omega^+} |X|\,d|\phi| + \int_{B\cap\Omega^+} -|X|\,d|\phi| - \int_{A\cap\Omega^-} |X|\,d|\phi| - \int_{B\cap\Omega^-} -|X|\,d|\phi|\right| \\
&= \left|\int_{A\cap\Omega^+} |X|\,d|\phi| - \int_{B\cap\Omega^+} |X|\,d|\phi| - \int_{A\cap\Omega^-} |X|\,d|\phi| + \int_{B\cap\Omega^-} |X|\,d|\phi|\right| \\
&\leq \left|\int_{A\cap\Omega^+} |X|\,d|\phi| + \int_{B\cap\Omega^+} |X|\,d|\phi| + \int_{A\cap\Omega^-} |X|\,d|\phi| + \int_{B\cap\Omega^-} |X|\,d|\phi|\right| \\
&= \int_{A\cap\Omega^+} |X|\,d|\phi| + \int_{B\cap\Omega^+} |X|\,d|\phi| + \int_{A\cap\Omega^-} |X|\,d|\phi| + \int_{B\cap\Omega^-} |X|\,d|\phi| \\
&= \int_\Omega |X|\,d|\phi|
\end{aligned}$$

Problem 5

If we have random variables $U \sim \text{Unif}(0,1)$ and $P \sim \text{Poisson}(\lambda)$, the the distribution functions are:

$$F_U(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases} \quad \text{and} \quad F_P(x) = \begin{cases} 0 & x < 0 \\ e^{-\lambda}\displaystyle\sum_{i=0}^{\lfloor x\rfloor}\frac{\lambda^i}{i!} & x \geq 0 \end{cases}$$

Now if we define a random variable X to be U if a coin is heads and P if a coin is tails, then $P(X \le x) = \frac{1}{2}P(U \le x) + \frac{1}{2}P(P \le x)$ or in other words $F_X(x) = \frac{1}{2}F_U(x) + \frac{1}{2}F_P(x)$ so explicitly we have:

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x + e^{-\lambda}}{2} & 0 \le x < 1 \\ \dfrac{1}{2} + \dfrac{e^{-\lambda}}{2} \displaystyle\sum_{i=0}^{\lfloor x \rfloor} \dfrac{\lambda^i}{i!} & x \ge 1 \end{cases}$$

Finally we define the corresponding measure $\phi((a,b]) = F_X(b) - F_X(a)$ (that we then extend with Carathéodory and then complete).

(a) We want to find the Lebesgue decomposition $\phi = \phi_{\mathrm{ac}} + \phi_{\mathrm{s}}$ w.r.t. Lebesgue measure λ. First consider this example — define the decreasing sequence of intervals $I_k = (1 - \frac{1}{k}, 1]$ and so $\lim\limits_{k\to\infty} I_k = \{1\}$. Then as long as the measure of the sets in question are not infinite, limits of of decreasing sets and measures commute, so

$$\lambda\left(\lim_{k\to\infty} I_k\right) = \lim_{k\to\infty} \frac{1}{k} = 0 \quad \text{and} \quad \phi\left(\lim_{k\to\infty} I_k\right) = \phi(\{1\}) = \lim_{k\to\infty} F(1) - F(1 - \tfrac{1}{k}) = \frac{\lambda e^{-\lambda}}{2} \neq 0$$

Based on this, we see that in order to construct $\phi_{\mathrm{ac}} \ll \lambda$, we should avoid all the jump points. Additionally, we see that

- for $n \in \{0, 1, 2, \dots\} = \mathbb{Z}_{\ge 0}, \phi(\{n\}) = \dfrac{e^{-\lambda}\lambda^n}{2(n!)}$ (while $\lambda(\{n\}) = 0$ sadly)
- for $(a,b] \subseteq (0,1), \phi((a,b]) = \dfrac{b-a}{2} \le b - a = \lambda((a,b])$
- for any intervals J s.t. $J \cap (\mathbb{Z}_{\ge 0} \cup (0,1)) = \varnothing$, we have that $\phi(J) = 0$ (obviously $\le$ whatever $\lambda(J)$ is).

With all this in mind, set $\phi_{\mathrm{ac}}(A) = \phi\big(A \setminus \mathbb{Z}_{\ge 0}\big)$ (justification from the above bullet points) and $\phi_{\mathrm{s}}(A) = \phi\big(A \cap \mathbb{Z}_{\ge 0}\big)$ (where we can let $N = \mathbb{Z}_{\ge 0}$ so that $\lambda(N) = 0$, and $\phi_{\mathrm{s}}(N^{\complement}) = \phi(\mathbb{Z}_{\ge 0}^{\complement} \cap \mathbb{Z}_{\ge 0}) = \phi(\varnothing) = 0$).

(b) Now considering the counting measure $\#$ over $\mathbb{Z}_{\ge 0}$; if $\#(A) = 0$ for some set, then $A = \varnothing$ because if A had just one element from $\mathbb{Z}_{\ge 0}$, the counting measure would return something greater than 0. $\phi(\varnothing) = 0$ obviously, so $\phi \ll \#$ so $\phi_{\mathrm{ac}} = \phi$.

Problem 6

We have a bounded, increasing, right-continuous function F on $\mathbb{R}$ where $F(-\infty) = 0$. Define $\mu_F((a,b]) = F(b) - F(a)$. We want to prove that $\mu_F \ll \lambda$ (Lebesgue measure) $\iff$ μ_F is absolutely continuous, where absolutely continuous (on $\mathbb{R}$) is defined as for all $\epsilon > 0$, there is δ_ϵ s.t.

$$\sum_{k=1}^{n} (b_k - a_k) < \delta_\epsilon \implies \sum_{k=1}^{n} |F(b_k) - F(a_k)| < \epsilon$$

for any finite sequence of disjoint subintervals $(a_k, b_k]$. Because F is increasing, we can get rid of the absolute values. Furthermore, we can rephrase this definition of absolute continuity to be that for every $\epsilon > 0$, there is δ_ϵ such that

$$\lambda(A) < \delta_\epsilon \implies \mu_F(A) < \epsilon \quad \text{for any } A = \bigcup_{k=1}^{n} (a_k, b_k].$$

So our problem is equivalent to proving $\mu_F \ll \lambda \iff [\lambda(A) < \delta_\epsilon \implies \mu_F(A) < \epsilon]$ (where the sets A are described above).

We will actually prove a more general result first: if μ, ν are measures, and μ is finite, then

$$\mu \ll \nu \iff [\forall \epsilon > 0, \exists \delta > 0 \text{ s.t. } \nu(A) < \delta_\epsilon \implies \mu(A) < \epsilon] \quad (\text{where } A \in \mathcal{A})$$

($\implies$) Suppose not. Then, $\exists \epsilon > 0$ s.t. $\forall \delta > 0$, $\exists A \in \mathcal{A}$ s.t. $\mu(A) \geq \epsilon$ and $\nu(A) < \delta$. More specifically, $\forall k \in \mathbb{N}, \exists A_k \in \mathcal{A}$ s.t. $\nu(A_k) < \dfrac{1}{2^k}$ and $\mu(A_k) \geq \epsilon$. Define the decreasing sequence of sets $B_k = \bigcup\limits_{i=k}^{\infty} A_i$. From that definition, we can say that

$$\nu(B_k) \leq \sum_{i=k}^{\infty} \nu(A_i) < \frac{2}{2^k} = \frac{1}{2^{k-1}}.$$

Then, define $B = \lim\limits_{k\to\infty} B_k = \bigcap\limits_{k=1}^{\infty} B_k$. Then because limits of monotone decreasing sets and measures commute (again, as long as the measure of the sets in the sequence are not infinite), we know that

$$\nu(B) = \nu\left(\lim_{k\to\infty} B_k\right) = \lim_{k\to\infty} \nu(B_k) = \lim_{k\to\infty} \frac{1}{2^{k-1}} = 0$$

However, $\mu(B_k) \geq \mu(A_k) \geq \epsilon$, so by limit-measure commutativity (and that μ is finite), we know that

$$\mu(B) = \mu\left(\lim_{k\to\infty} B_k\right) = \lim_{k\to\infty} \mu(B_k) \geq \epsilon$$

But by $\ll$, $\nu(B) = 0 \implies \mu(B) = 0$; hence, contradiction.

($\impliedby$) This direction is quite simple: assume $\nu(N) = 0$ for some $N \in \mathcal{A}$. Then, we know that $\forall \epsilon > 0, \exists \delta > 0$ s.t. $\nu(A) < \delta \implies \mu(A) < \epsilon$. Well, $\nu(N) = 0$ which is less than δ for any δ, so $\forall \epsilon > 0, \mu(N) < \epsilon \implies \mu(N) = 0$ and so $\mu \ll \nu$.

Finally, we are given that F is bounded, so μ_F is finite. We can use the Carathéodory extension theorem and completion to extend μ_F (only defined for $A = \bigcup\limits_{k=1}^{n} (a_k, b_k]$, i.e. a field) to an actual measure (also finite), denoted $\overline{\mu_F}$. Then the general lemma above proves the result for all $A \in \mathcal{A}$, and of course in particular A of the aforementioned form (because the extension agrees with the pre-measure on the field). QED!

521 Homework 6

Daniel Rui - 11/6/19

Problem 1

(a) Consider the probability space $\left([0,1], \mathcal{B}_{[0,1]}, P\right)$ and the random variables $X_n = \dfrac{1}{nx}$. Then, $X_n \to_{\text{a.s.}} 0$. But $\mathbb{E}[X_n] = \infty$ for all $n \in \mathbb{N}$ whereas $\mathbb{E}[0] = 0$.

(b) Again on the probability space $\left([0,1], \mathcal{B}_{[0,1]}, P\right)$, take $X_n = \left(1 + \frac{1}{n}\right) 1_{[0,\frac{1}{2}]}$ and $X = 1_{[\frac{1}{2},1]}$. Then $\mathbb{E}[X_n] \to \frac{1}{2} = \mathbb{E}[X]$ but obviously $X_n \not\to_{\text{a.s.}} X$.

(c) Again on $\left([0,1], \mathcal{B}_{[0,1]}, P\right)$, let's define $X_n(\omega) = \frac{1}{n\omega} + 1 - \omega$ and $X' = \omega$. Denote $X = 1 - \omega$. Then $X_n \to_p X$. Because $X_n \to_p X \implies X_n \to_d X$ (from the midterm!), and we know that both X and X' have distribution function: 0 (if $x \leq 0$), x (if $0 < x < 1$), 1 (if $x \geq 1$). Thus, $X_n \to_d X'$.

However, $X_n \not\to_p X'$ (obviously), and $X_n \not\to_{\text{a.s.}} X'$ (also obviously), and $\mathbb{E}[X_n] = \infty$ while $\mathbb{E}[X'] = \frac{1}{2}$, so $X_n \not\to_{r=1} X'$.

Problem 2

Consider functions $f, f_1, f_2, \ldots \geq 0$.

(a) For this part, we have that $f_n \to_{\text{a.e.}} f$ and additionally that $\int_\Omega f_n \, d\mu = 1$ and $\int_\Omega f \, d\mu = 1$. We want to prove that $\sup\limits_{A \in \mathcal{A}} \left| \int_A f_n \, d\mu - \int_A f \, d\mu \right| \to 0$. Define the sequences $m_n = \min\{f, f_n\}$ and $M_n = \max\{f, f_n\}$. Because $f_n \to_{\text{a.e.}} f$, we know that $m_n \to_{\text{a.e.}} f$ and $M_n \to_{\text{a.e.}} f$.

Now note that $m_n \leq f$ (everywhere), where f is clearly integrable (with integral 1). So by the DCT,

$$\lim_{n \to \infty} \int_\Omega m_n \, d\mu = \int_\Omega f \, d\mu = 1$$

Furthermore, we see that because $m_n + M_n = f + f_n$,

$$\int_\Omega M_n \, d\mu = \int_\Omega f + f_n - m_n \, d\mu = \int_\Omega f \, d\mu + \int_\Omega f_n \, d\mu - \int_\Omega m_n \, d\mu = 1 + 1 - \int_\Omega m_n \, d\mu$$

and thus the limit as $n \to \infty$ of $\int_\Omega M_n \, d\mu$ is $1 + 1 - 1 = 1$. Thus,

$$\int_\Omega |f - f_n| \, d\mu = \int_\Omega M_n - m_n \, d\mu = \int_\Omega M_n \, d\mu - \int_\Omega m_n \, d\mu$$

which goes to $1 - 1 = 0$ as $n \to \infty$. Finally, for any $A \in \mathcal{A}$,

$$\left| \int_A f \, d\mu - \int_A f_n \, d\mu \right| \leq \int_A |f - f_n| \, d\mu \leq \int_\Omega |f - f_n| \, d\mu \to 0$$

as desired.

The conclusion above in fact holds if we just assume that $f_n \to_{\text{a.e.}} f$ and that $\int_\Omega f_n \, d\mu \to \int_\Omega f \, d\mu$: we again have that

$$\lim_{n\to\infty} \int_\Omega m_n \, d\mu = \int_\Omega f \, d\mu$$

and that

$$\int_\Omega M_n \, d\mu = \int_\Omega f \, d\mu + \int_\Omega f_n \, d\mu - \int_\Omega m_n \, d\mu$$

and so

$$\lim_{n\to\infty} \int_\Omega M_n \, d\mu = \int_\Omega f \, d\mu + \int_\Omega f \, d\mu - \int_\Omega f \, d\mu = \int_\Omega f \, d\mu$$

and lastly

$$\left| \int_A f \, d\mu - \int_A f_n \, d\mu \right| \le \int_A |f - f_n| \, d\mu \le \int_\Omega |f - f_n| \, d\mu = \int_\Omega M_n \, d\mu - \int_\Omega m_n \, d\mu \to 0$$

(b) For this part, we are given that $f_n \to_\mu f$ and that $\int_\Omega f_n \, d\mu \to \int_\Omega f \, d\mu$. Because the f_n converge in measure to f, we know that for every subsequence f_{n_k}, we can find a subsubsequence $f_{n_{k_i}}$ that converges pointwise almost everywhere to f. Convergence a.e. and convergence of the integrals over Ω means that we can use the general result from the end of part (a):

$$\left| \int_A f \, d\mu - \int_A f_{n_{k_i}} \, d\mu \right| \le \int_A |f - f_{n_{k_i}}| \, d\mu \le \int_\Omega |f - f_{n_{k_i}}| \, d\mu \to 0.$$

Denote the sequence of real numbers $a_n = \int_A f_n \, d\mu$. Then, we just proved that for every subsequence a_{n_k}, the subsubsequence $a_{n_{k_i}}$ converges to $a = \int_A f \, d\mu$. Given only this, it turns out that a_n *must* converge to a. We prove this by contradiction: suppose that a_n does not converge to a. That means that for every $\epsilon > 0$ and $k \in \mathbb{N}$, we could find $n_k > k$ s.t. $|a_{n_k} - a| \ge \epsilon$. Then by construction, the subsequence a_{n_k} *would not* have a convergent subsubsequence; contradiction. Thus, $a_n \to a$, or in other words, for all $A \in \mathcal{A}$,

$$\left| \int_A f \, d\mu - \int_A f_n \, d\mu \right| = |a - a_n| \to 0.$$

Because this holds for all $A \in \mathcal{A}$, it holds for the supremum and so the result follows.

521 Homework 5

Daniel Rui - 10/30/19

Problem 1

We define Pearson's correlation coefficient (where $\mu_X = \mathbb{E}[X]$) as

$$\rho = \text{Corr}[X,Y] = \frac{\text{Cov}[X,Y]}{\sqrt{\text{Var}[X]\text{Var}[Y]}} = \frac{\mathbb{E}[(X-\mu_X)(Y-\mu_Y)]}{\sqrt{\mathbb{E}[(X-\mu_X)^2]\mathbb{E}[(Y-\mu_Y)^2]}}$$

We have that

$$\begin{aligned} \rho = \pm 1 &\iff \mathbb{E}[(X-\mu_X)(Y-\mu_Y)] = \pm\sqrt{\mathbb{E}[(X-\mu_X)^2]\mathbb{E}[(Y-\mu_Y)^2]} \\ &\iff \Big(\mathbb{E}[(X-\mu_X)(Y-\mu_Y)]\Big)^2 = \mathbb{E}[(X-\mu_X)^2]\mathbb{E}[(Y-\mu_Y)^2] \end{aligned}$$

This is the equality case of the Cauchy-Schwarz inequality, which we know is satisfied if and only if $(Y-\mu_Y) = a(X-\mu_X)$ a.e. for some $a \neq 0$. (Note that all the above implications are double sided, so everything so far is "if and only if")

Multiplying both sides by the random variable $(X-\mu_X)$, we get $(X-\mu_X)(Y-\mu_Y) = a(X-\mu_X)^2$ a.e. $\iff \mathbb{E}[(X-\mu_X)(Y-\mu_Y)] = a\mathbb{E}[(X-\mu_X)^2]$. $\mathbb{E}[(X-\mu_X)^2] > 0$, so $a > 0 \iff \mathbb{E}[(X-\mu_X)(Y-\mu_Y)]$ (the numerator of ρ) is positive $\iff$ ρ must be positive and hence equal to 1. Similarly, if $a < 0 \iff \rho$ must be negative and hence must equal -1.

Problem 2

Let $\mu_r = \mathbb{E}[|X|^r]$. We want to show that for $r \geq s \geq t \geq 0$, $\mu_r^{s-t}\mu_t^{r-s} \geq \mu_s^{r-t}$, or more transparently,

$$\Big(\mathbb{E}[|X|^r]\Big)^{s-t} \cdot \Big(\mathbb{E}[|X|^t]\Big)^{r-s} \geq \Big(\mathbb{E}[|X|^s]\Big)^{r-t}$$

Hölder's inequality gives that for $X_1, X_2 \geq 0$ and $a, b > 1$ s.t. $\frac{1}{a} + \frac{1}{b} = 1$,

$$\mathbb{E}[X_1 X_2] \leq \Big(\mathbb{E}[X_1^a]\Big)^{1/a}\Big(\mathbb{E}[X_2^b]\Big)^{1/b}$$

To fit our problem to something Hölder's inequality could help with, we definitely want the outer exponents to match, so we want that

$$\frac{1}{a} = \frac{s-t}{r-t} \quad \text{and} \quad \frac{1}{b} = \frac{r-s}{r-t}$$

Note that $s - t \leq r - t$. In the case that $r > s$, then $\frac{1}{a} < 1 \implies a > 1$ so we can use Hölder's inequality. In the case that $r = s$, then the problem statement is trivial: $\mu_r^{s-t} = \mu_s^{r-t}$. So for further analysis, let's just focus on $a > 1$ (and similarly $b > 1$).

Now given these values for a, b, we now want the inner exponents to match by setting $X_1 = |X^c| = |X|^c$ and $X_2 = |X^d| = |X|^d$:

$$|X|^r = X_1^a = |X|^{ca} \implies c = \frac{r}{a} = \frac{r(s-t)}{r-t} \quad \text{and} \quad |X|^t = X_1^b = |X|^{db} \implies d = \frac{t}{b} = \frac{t(r-s)}{r-t}$$

Very fortunately, we see that $c + d = \dfrac{r(s-t) + t(r-s)}{r-t} = \dfrac{rs - st}{r-t} = s$, so $X_1 X_2 = |X|^{c+d} = |X|^s$.

Thus, setting a, b, c, d, X_1, X_2 as described gives that Hölder's inequality is equivalent to the desired result.

Problem 3

Define random i.i.d. variables $\epsilon_1, \ldots, \epsilon_n$ with $P(\epsilon_i = \pm 1) = \frac{1}{2}$, and $a, b, a_i \in \mathbb{R}$. Note that $\mathbb{E}[\epsilon_i] = 1 \cdot \frac{1}{2} + (-1) \cdot \frac{1}{2} = 0$ for all i. We want to prove one particular case of Khintchine's inequality (the case where $p = 1$):

$$a\sqrt{\sum_{i=1}^n a_i^2} \leq \mathbb{E}\left[\left|\sum_{i=1}^n a_i \epsilon_i\right|\right] \leq b\sqrt{\sum_{i=1}^n a_i^2}$$

The right inequality is relatively straightforward: using Minkowski's inequality (with $r = 1$) and Cauchy-Schwarz, we see that

$$\mathbb{E}\left[\left|\sum_{i=1}^n a_i \epsilon_i\right|\right] \leq \sum_{i=1}^n \mathbb{E}[|a_i \epsilon_i|] = \sum_{i=1}^n \mathbb{E}[|a_i||\epsilon_i|] \leq \sqrt{\sum_{i=1}^n \mathbb{E}[|a_i|^2]\mathbb{E}[|\epsilon_i|^2]} = \sqrt{\sum_{i=1}^n a_1^2}$$

The left inequality is much more non-trivial: we first observe two facts. Define Z to be the random variable $\sum\limits_{i=1}^n a_i \epsilon_i$. Mini-lemma: $\mathbb{E}[Z^2] = \sum\limits_{i=1}^n a_i^2$, proof as follows:

$$\mathbb{E}[Z^2] = \mathbb{E}\left[\sum_{i=1}^n \sum_{j=1}^n a_i a_j \epsilon_i \epsilon_j\right] = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \mathbb{E}[\epsilon_i \epsilon_j] = \sum_{i=1}^n a_i^2$$

where the last equality holds because if $i \neq j$, then ϵ_i and ϵ_j are independent so $\mathbb{E}[\epsilon_i \epsilon_j] = \mathbb{E}[\epsilon_i]\mathbb{E}[\epsilon_j] = 0$, and if $i = j$ then $\mathbb{E}[\epsilon_i \epsilon_j] = \mathbb{E}[\epsilon_i^2] = 1^2 \cdot \frac{1}{2} + (-1)^2 \cdot \frac{1}{2} = 1$.

Our second fact is more substantial, so lemma: $\mathbb{E}[Z^4] \leq 3(\mathbb{E}[Z^2])^2 = 3\left(\sum\limits_{i=1}^n a_i^2\right)^2$. We prove this by induction. For the base case, $n = 2$,

$$\begin{aligned}
\mathbb{E}[(a_1\epsilon_1 + a_2\epsilon_2)^4] &= \mathbb{E}[a_1^4\epsilon_1^4 + 4a_1^3\epsilon_1^3 a_2\epsilon_2 + 6a_1^2\epsilon_1^2 a_2^2\epsilon_2^2 + 4a_1\epsilon_1 a_2^3\epsilon_2^3 + a_2^4\epsilon_2^4] \\
&= a_1^4\mathbb{E}[\epsilon_1^4] + 4a_1^3 a_2\mathbb{E}[\epsilon_1^3\epsilon_2] + 6a_1^2 a_2^2\mathbb{E}[\epsilon_1^2\epsilon_2^2] + 4a_1 a_2^3\mathbb{E}[\epsilon_1\epsilon_2^3] + a_2^4\mathbb{E}[\epsilon_2^4] \\
&= a_1^4 + 6a_1^2 a_2^2 + a_2^4 \\
&\leq 3a_1^4 + 6a_1^2 a_2^2 + 3a_2^4 = 3(a_1^2 + a_2^2)^2
\end{aligned}$$

Now assume that the inequality holds for $n = k-1$. Denote $S = a_1\epsilon_1 + \ldots + a_{k-1}\epsilon_{k-1}$. Then

$$\begin{aligned}
\mathbb{E}[(a_1\epsilon_1 + \ldots + a_k\epsilon_k)^4] &= \mathbb{E}[(S + a_k\epsilon_k)^4] \\
&= \mathbb{E}[S^4 + 4S^3 a_k\epsilon_k + 6S^2 a_k^2\epsilon_k^2 + 4Sa_k^3\epsilon_k^3 + a_k^4\epsilon_k^4] \\
&= \mathbb{E}[S^4] + 4a_k\mathbb{E}[S^3]\mathbb{E}[\epsilon_k] + 6a_k^2\mathbb{E}[S^2]\mathbb{E}[\epsilon_k^2] + 4a_k^3\mathbb{E}[S]\mathbb{E}[\epsilon_k^3] + a_k^4\mathbb{E}[\epsilon_k^4] \\
&= \mathbb{E}[S^4] + 6a_k^2\mathbb{E}[S^2] + a_k^4 \\
&\leq 3(a_1^2 + \ldots + a_{k-1}^2)^2 + 6a_k(a^2 + \ldots + a_{k-1}^2) + a_k^4 \\
&\leq 3(a_1^2 + \ldots + a_{k-1}^2)^2 + 6a_k(a^2 + \ldots + a_{k-1}^2) + 3a_k^4 \\
&= 3((a_1^2 + \ldots + a_{k-1}^2) + a_k^2)^2
\end{aligned}$$

Our final piece of information will be derived from Problem 2; setting $(r,s,t) = (4,2,1)$, we have that

$$\Big(\mathbb{E}[Z^2]\Big)^3 \leq \Big(\mathbb{E}[|Z|]\Big)^2 \mathbb{E}[Z^4] \implies \mathbb{E}[|Z|] \geq \sqrt{\frac{(\mathbb{E}[Z^2])^3}{\mathbb{E}[Z^4]}}$$

Using our lemma, $\mathbb{E}[Z^4] \leq 3(\mathbb{E}[Z^2])^2 \implies \dfrac{1}{\mathbb{E}[Z^4]} \geq \dfrac{1}{3(\mathbb{E}[Z^2])^2}$, on this most recent inequality, we get

$$\mathbb{E}[|Z|] \geq \sqrt{\frac{(\mathbb{E}[Z^2])^3}{\mathbb{E}[Z^4]}} \geq \sqrt{\frac{(\mathbb{E}[Z^2])^3}{3(\mathbb{E}[Z^2])^2}} = \sqrt{\frac{\mathbb{E}[Z^2]}{3}} = \frac{\sqrt{3}}{3}\sqrt{\sum_{i=1}^n a_i^2}$$

and so in conclusion

$$\frac{\sqrt{3}}{3} \cdot \sqrt{\sum_{i=1}^n a_i^2} \leq \mathbb{E}\left[\left|\sum_{i=1}^n a_i\epsilon_i\right|\right] \leq 1 \cdot \sqrt{\sum_{i=1}^n a_i^2}$$

Problem 4

(a) Consider a random variable $X \geq 0$, which has a distribution function $F(x) = P([X < x])$. We want to show that

$$\mathbb{E}[X] = \int_\Omega X\,dP = \int_0^\infty P([X \geq x])\,dx$$

Of course, we proceed using the standard method of starting with indicator functions and working upwards.

(i) For a set $A \in \mathcal{A}$, take $X = 1_A$ (notice that X never exceeds 1, and that $[X = 1] = A$ and $[X = 0] = \Omega \setminus A$). Then,

$$\int_\Omega X\,dP = 1 \cdot P(A) \quad \text{and} \quad \int_0^\infty P([X \geq x])\,dx = \int_0^1 P([X \geq x])\,dx = \int_0^1 P(A)\,dx = P(A)$$

(ii) Now we look at simple functions: $X = \sum_{i=1}^n c_i 1_{A_i}$ for (pairwise) disjoint A_i and $c_i \in \mathbb{R}$ such that $c_i \neq c_j$ when $i \neq j$. Furthermore, let's order the terms so that the c_i are increasing.

Then,

$$\int_\Omega X\,dP = \sum_{i=1}^n c_i P(A_i)$$

and $\displaystyle\int_0^\infty P([X \geq x])\,dx$

$$\begin{aligned}
&= \int_{c_{n-1}}^{c_n} P(A_n)\,dx + \int_{c_{n-2}}^{c_{n-1}} P(A_{n-1} \cup A_n)\,dx + \ldots + \int_0^{c_1} P(A_1 \cup \ldots \cup A_n)\,dx \\
&= \int_{c_{n-1}}^{c_n} P(A_n)\,dx + \ldots + \int_0^{c_1} P(A_1) + \ldots + P(A_n)\,dx \\
&= \int_0^{c_n} P(A_n)\,dx + \int_0^{c_{n-1}} P(A_{n-1})\,dx + \ldots \int_0^{c_1} P(A_1)\,dx \\
&= c_n P(A_n) + c_{n-1}P(A_{n-1}) + \ldots c_1 P(A_1) = \sum_{i=1}^n c_i P(A_i)
\end{aligned}$$

(iii) Finally, we know that for any random variable $X \geq 0$, $X_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \cdot \left(1_{X^{-1}\left(\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)\right)}\right) + n \cdot \left(1_{\{\omega : X \geq n\}}\right)$ is a monotone increasing sequence converging pointwise everywhere to X. Additionally, because the X_n are increasing, $[X_n \geq x] \subseteq [X_{n+1} \geq x]$, so the $P([X_n \geq x])$ are also monotone increasing to $P([X \geq x])$ (?). From part (ii), we know that

$$\int_\Omega X_n\,dP = \int_0^\infty P([X_n \geq x])\,dx$$

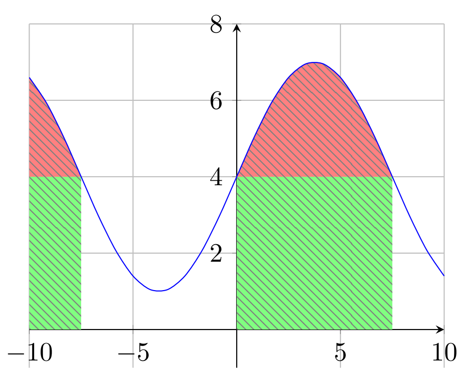
for any $n \in \mathbb{N}$, and so by the monotone convergence theorem,

$$\lim_{n\to\infty} \int_\Omega X_n\,dP = \lim_{n\to\infty} \int_0^\infty P([X_n \geq x])\,dx \implies \int_\Omega X\,dP = \int_0^\infty P([X \geq x])\,dx$$

Last word: $1 - F(x) = P([X \geq x])$ so we can put $1 - F(x)$ in the above integrals.

(b)

$$\int_{[X \geq \lambda]} X\,dP = \lambda P([X \geq \lambda]) + \int_\lambda^\infty P([X \geq x])\,dx$$



The proof is as follows: $\int_{[X\geq\lambda]} X\,dP$ is equivalent to $\int_\Omega 1_{[X\geq\lambda]}\cdot X\,dP$. Denote $X' = 1_{[X\geq\lambda]}\cdot X$, and observe that $P([X'\geq x]) = P([X\geq\lambda])$ for all $x\in(0,\lambda]$ (by definition of X'). By part (a) we know that

$$\begin{aligned}
\int_{[X\geq\lambda]} X\,dP &= \int_\Omega X'\,dP \\
&= \int_0^\lambda P([X'\geq x])\,dx + \int_\lambda^\infty P([X'\geq x])\,dx \\
&= \int_0^\lambda P([X\geq\lambda])\,dx + \int_\lambda^\infty P([X\geq x])\,dx \\
&= \lambda P([X\geq\lambda]) + \int_\lambda^\infty P([X\geq x])\,dx
\end{aligned}$$

521 Homework 4

Daniel Rui - 10/23/19

Problem 1

Important theorem from class/the book: if $\mu(\Omega) < \infty$, $X_n \to_\mu X \iff$ for every subsequence $X_{n'}$, there is a further subsequence $X_{n''}$ that converges a.e. to X.

(a) We are given that $\mu(\Omega) < \infty$ and that g is continuous a.e. w.r.t μ_X, i.e. g is continuous on a set $S^\complement$ where $\mu_X(S) = \mu(X^{-1}(S)) = \mu([X \in S]) = 0$, and that $X_n \to_\mu X$. Consider now the set N_1 of all ω_0 s.t. we can always find ω arbitrarily close to ω_0 s.t. $|g(X(\omega)) - g(X(\omega_0))| > \epsilon$ for some $\epsilon > 0$. Because S is the set of all points where g is discontinuous, $X(N_1) \subseteq S$. In other words, N_1 is *precisely* the set of all points that X maps to S, or $[X \in S]$. But we know this set has measure zero, so $\mu(N_1) = 0$.

Our final piece of information is that $X_n \to_\mu X$, which we know from the theorem above to mean that for every subsequence $X_{n'}$, we can find a subsubsequence $X_{n''} \to_{\text{a.e.}} X$, i.e. $X_{n''}$ converges to X on a set $N_2^\complement$ where $\mu(N_2) = 0$.

Now for all $\omega \in N_1^\complement \cap N_2^\complement$, we see that $X_{n''}(\omega) \to X(\omega)$ (pointwise convergence) and that g is continuous. From the definitions of convergence and continuity, we know that $\forall \epsilon > 0, \exists \delta > 0, \exists \epsilon' < \delta$ and $\exists N \in \mathbb{N}$ s.t. for all $n'' \geq N$,

$$|X_{n''}(\omega) - X(\omega)| < \epsilon' < \delta \implies |g(X_{n''}(\omega)) - g(X(\omega))| < \epsilon$$

This is true for all $\omega \in (N_1 \cup N_2)^\complement$, so $g(X_{n''}(\omega))$ does not converge to $g(X(\omega))$ *only* on $N_1 \cup N_2$. But $\mu(N_1 \cup N_2) \leq \mu(N_1) + \mu(N_2) = 0$, so in fact $g \circ X_{n''} \to_{\text{a.e.}} g \circ X$. We can find such a subsubsequence for all subsequences $X_{n'}$, so $g \circ X_n \to_\mu g \circ X$.

(b) For part (b), we allow $\mu(\Omega) = \infty$. g is now uniformly continuous on $\mathbb{R}$, and $X_n \to_\mu X$. From uniformly continuity, for every ϵ we can find one δ_ϵ s.t. for all x, y on the real line, $|x - y| < \delta_\epsilon \implies |g(x) - g(y)| < \epsilon$. By definition, $X_n \to_\mu X \iff$ for all $\epsilon' > 0$, we can find arbitrarily small δ' such that for n beyond some $N_{\epsilon',\delta'}$, $\mu(A_{\epsilon'}) < \epsilon'$, where $A_{\epsilon'} = \{\omega \in \Omega : |X_n(\omega) - X(\omega)| \geq \delta'\}$. Note that $A_{\epsilon'}^\complement = \{\omega \in \Omega : |X_n(\omega) - X(\omega)| < \delta'\}$ (for $n \geq N_{\epsilon',\delta'}$ and arbitrarily small ϵ' and δ').

Because δ' is arbitrarily small, we can always let $\delta' < \delta_\epsilon$, so by uniform continuity, $|X_n(\omega) - X(\omega)| < \delta' < \delta_\epsilon \implies |g(X_n(\omega)) - g(X(\omega))| < \epsilon$ (again only for $n \geq N_{\epsilon',\delta'}$). In other words, $A_{\epsilon'}^\complement \subseteq B_\epsilon^\complement$ where $B_\epsilon^\complement = \{\omega \in \Omega : |g(X_n(\omega)) - g(X(\omega))| < \epsilon\}$. Thus, $B_\epsilon \subseteq A_{\epsilon'}$ so $\mu(B_\epsilon) < \epsilon'$.

Tying everything together, we just proved that for all $\epsilon' > 0$, we can find an arbitrarily small ϵ s.t. if $n \geq N_{\epsilon',\delta'}$ (the exact same $N_{\epsilon',\delta'}$ we found from the definition of $X_n \to_\mu X$ which we know to exist), then $\mu(B_\epsilon) < \epsilon'$. Hence, by definition, $g \circ X_n \to_\mu g \circ X$.

Problem 2

We are given a measurable function $X \geq 0$ and that $\displaystyle\int_\Omega X\,d\mu = 0$. Now define a sequence of sets $\{A_n\}$ where $A_n = [X > \frac{1}{n}] = \{\omega \in \Omega : X(\omega) > \frac{1}{n}\}$. This sequence satisfies $A_n \subseteq A_{n+1}$, so it is monotonically increasing. Now note that

$$0 = \int_\Omega X\,d\mu \geq \int_{A_n} X\,d\mu \geq \int_{A_n} \frac{1}{n}\,d\mu = \frac{1}{n} \cdot \mu(A_n) \geq 0$$

which means that $\mu(A_n) = 0$ for all of our sets A_n. Then because A_n is increasing we get that

$$\mu\left([X > 0]\right) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} \mu\left(A_n\right) = 0$$

Problem 3

We are given an arbitrary measurable function X that satisfies $\displaystyle\int_A X\,d\mu = 0$ for all $A \in \mathcal{A}$.

(a) Define the increasing sequence of sets $\{A_n\}$ where $A_n = [X < -\frac{1}{n}] \cup [X > \frac{1}{n}]$. Because X is $\mathcal{A}$-measurable, all the $A_n = X^{-1}(B)$ — where B is the Borel set $(-\infty, -\frac{1}{n}) \cup (\frac{1}{n}, \infty)$ — are in $\mathcal{A}$, so by the given, $\int_{A_n} X\,d\mu = 0$ for all A_n. Using the same argument as from Problem 2, we see that $\mu(A_n) = 0$ for all A_n and hence again $\mu(\bigcup_{n=1}^{\infty}) = \mu([X \neq 0]) = 0$, i.e. $X \neq 0$ only on a set of measure zero, so $X = 0$ a.e.

(b) Define the increasing sequence of sets $\{A_n\}$ where $A_n = [X < -\frac{1}{n}]$. Similar to part (a), we note that $\int_{A_n} X\,d\mu \geq 0$ for all A_n. But now

$$0 \leq \int_{A_n} X\,d\mu \leq \int_{A_n} -\frac{1}{n}\,d\mu \leq -\frac{1}{n}\mu(A_n)$$

but μ can never be negative, so $-\frac{1}{n}\mu(A_n) \leq 0$. We've sandwiched $-\frac{1}{n}\mu(A_n)$ between 0 and 0, so $\mu(A_n) = 0$ for all A_n. Thus, like above, we can say that $\mu([X < 0]) = 0$, i.e. $X \geq 0$ a.e.

Problem 4

The theorem of the unconscious statistician:

$$\int_{X^{-1}(g^{-1}(B))} g(X(\omega))\,d\mu(\omega) = \int_{g^{-1}(B)} g(x)\,d\mu_X(x) = \int_B y\,d\mu_Y(y)$$

where $X : (\Omega, \mathcal{A}, \mu) \to (\Omega', \mathcal{A}', \mu_X)$; $\mu_X(A') = \mu(X^{-1}(A')) = \mu([X \in A'])$; and $g : (\Omega', \mathcal{A}') \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$; define $Y := g(X(\omega))$, and $\mu_Y(B) = \mu_{g\circ X}\big(B\big) = \mu\big((g \circ X)^{-1}(B)\big) = \mu\big(X^{-1} \circ g^{-1}(B)\big) = \mu_X\big(g^{-1}(B)\big)$.

The first equality is covered in the book. Like in the book, we split up the proof into four steps:

(i) Is y a function? In the preceding integrals ω, x were not ... :(

521 Homework 3

Daniel Rui - 10/16/19

Problem 1

We have measurable functions $X_1(x,y) = x$ and $X_2(x,y) = y$ (for all $(x,y) \in \mathbb{R}^2$), and $Z_1(x,y) = \sqrt{X_1^2 + X_2^2}$ and $Z_2(x,y) = \operatorname{sign}(X_1 - X_2)$. Define $\mathcal{F}(X) = X^{-1}(\bar{\mathcal{B}})$ (where $\bar{\mathcal{B}} = \sigma\Big[\mathcal{B}, \{\infty\}, \{-\infty\}\Big]$) to be $X^{-1}(\bar{\mathcal{B}}) = \{X^{-1}(B) : B \in \bar{\mathcal{B}}\} = \Big\{\{(x,y) \in \mathbb{R}^2 : X(x,y) \in B\} : B \in \bar{\mathcal{B}}\Big\}$ or in English, the set of all the sets of points that X sends to a Borel set, for all Borel sets.

(a) Consider $\mathcal{F}(Z_1) = Z_1^{-1}(\bar{\mathcal{B}}) = \Big\{\{(x,y) \in \mathbb{R}^2 : Z_1(x,y) \in B\} : B \in \bar{\mathcal{B}}\Big\}$. Let's start simply and consider what happens when $B = \{r\}$ for some $r \in \mathbb{R}$. If $r < 0$, then $Z_1^{-1}(B) = \varnothing$. For $r \geq 0$, $Z_1^{-1}(B)$ is a circle centered at $(0,0)$ with radius r. Now if B is some interval (r_1, r_2) with $r_2 > r_1 \geq 0$, then $Z_1^{-1}(B)$ is an open torus.

We can see that for whatever B is, $Z_1^{-1}(B)$ is some combination of circles, balls (filled in circles), and tori. But all of these can be generated by countable unions, intersections, and complements of plain old open balls, so $\mathcal{F}(Z_1)$ is the σ-algebra of the open balls in $\mathbb{R}^2$. Just for fun, a circle with radius r can be represented as $\bigcap\limits_{n=1}^{\infty} B\left(r + \frac{1}{n}\right) \setminus \bigcup\limits_{n=1}^{\infty} B\left(r - \frac{1}{n}\right)$ where $B(r)$ denotes the open ball centered at $(0,0)$ with radius r.

(b) $\mathcal{F}(Z_2) = \Big\{\{(x,y) \in \mathbb{R}^2 : Z_2(x,y) \in B\} : B \in \bar{\mathcal{B}}\Big\}$ is considerable easier — if $\{-1\} \in B$, then $Z_2^{-1}(B)$ will contain the lower right half of the plane, or $y < x$; if $\{0\} \in B$, then $Z_2^{-1}(B)$ will contain the line $y = x$; and if $\{1\} \in B$, then $Z_2^{-1}(B)$ will contain the upper left half of the plane $y > x$. Any Borel set B will either contain or not contain any one of these three points, so $\mathcal{F}(Z_2)$ will just be the set of all combinations of (a.k.a. the (σ-)algebra containing) $\{(x,y) \in \mathbb{R}^2 : y < x\}, \{(x,y) \in \mathbb{R}^2 : y = x\}$, and $\{(x,y) \in \mathbb{R}^2 : y > x\}$.

(c) The σ-algebra of the union of $\mathcal{F}(Z_1)$ and $\mathcal{F}(Z_2)$ is just the σ-algebra containing all half open balls (upper left and lower right) and all sets of two points on $y = x$ equidistant from the origin.

Problem 2

Let $\mathcal{C}$ be a $\bar{\pi}$-system of subsets of Ω (closed under finite intersections, and $\Omega \in \mathcal{C}$). Then let $\mathcal{V}$ be a vector space ($X, Y \in \mathcal{V} \implies X + Y \in \mathcal{V}$ and $\alpha X \in \mathcal{V}$) such that the characteristic function $1_C \in \mathcal{V}$ for all $C \in \mathcal{C}$ and that if $\{A_n\}$ is a sequence of monotonically increasing sets ($A_n \subseteq A_{n+1}$) s.t. $1_{A_n} \in \mathcal{V}$, then $A = \bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{V}$.

(a) For all $A \in \sigma[\mathcal{C}]$, $1_A \in \mathcal{V}$ — we prove this by proving that the set $\mathcal{S}$ of all sets S s.t. $1_S \in \mathcal{V}$ contains $\lambda[\mathcal{C}]$, which in conjunction with Dynkin's $\pi - \lambda$ theorem gives that $S \supseteq \lambda[\mathcal{C}] = \sigma[\mathcal{C}]$:

[Ω :] $\Omega \in \mathcal{S}$ obviously because $1_\Omega \in \mathcal{V}$; [$B \setminus A$:] if we are given that $A, B \in \mathcal{S}$ and $A \subset B$, then $1_B, 1_A \in \mathcal{V}$ so $1_{B \setminus A} = 1_B - 1_A \in \mathcal{V} \implies B \setminus A \in \mathcal{S}$; and [monotone unions:] if we are given that monotone increasing sequence of A_n is in $\mathcal{S}$, then $1_{A_n} \in \mathcal{V} \implies 1_{\cup A_n} \in V \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{S}$. Hence, $\mathcal{S}$ is a λ-system, as desired.

(b) By the axioms of vector spaces, for a partition A_i of Ω where $1_{A_i} \in \mathcal{V}$, the finite sum $\sum_{i=1}^{m} x_i 1_{A_i}$ will also be in $\mathcal{V}$.

(c) Some motivation for the nice construction in the book; for any measurable function $X(\omega)$ we are given (further assuming that $X \geq 0$ for all $\omega \in \Omega$, because for any general $X = X^+ - X^-$ can be represented as a linear combination of non-negative functions), let us define a sequence

$$X_n(\omega) = \sum_{k=1}^{n^2} \tfrac{k-1}{n} \cdot \left(1_{X^{-1}\left(\left[\frac{k-1}{n}, \frac{k}{n}\right)\right)}\right) + n \cdot \left(1_{\{\omega : X \geq n\}}\right)$$

For every $\omega \in \Omega$, there will be some N s.t. for all $n \geq N$, $X(\omega) < n$, which will imply that $|X_n(\omega) - X(\omega)| < \frac{1}{n}$, which means that $X_n \to X$. However, X_n does not increase monotonically (which we need), because $X_n(\omega)$ basically returns the largest multiple of $\frac{1}{n}$ less than or equal to $X(\omega)$, and the largest multiple of $\frac{1}{n} \leq X(\omega)$ may in fact be less than the largest multiple of $\frac{1}{m} \leq X(\omega)$ for $n > m$. Thus, we change our tactic a little bit to fix this error:

$$X_n(\omega) = \sum_{k=1}^{n2^n} \tfrac{k-1}{2^n} \cdot \left(1_{X^{-1}\left(\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)\right)}\right) + n \cdot \left(1_{\{\omega : X \geq n\}}\right)$$

It's easy to check that similar to above, for every $\omega \in \Omega$, there will be some N s.t. for all $n \geq N$, $X(\omega) < n$, which will imply that $|X_n(\omega) - X(\omega)| < 2^{-n}$. Now we have an monotone increasing sequence of X_n converging to X for any given ω, so by the given in the problem, $X \in \mathcal{V}$.

Interlude

Just a copy of the definitions of convergence almost everywhere:

$$\mu(\{\omega \in \Omega : \lim_{n\to\infty} |X_n(\omega) - X(\omega)| \neq 0\}) = 0$$

and convergence in measure

$$(\forall \epsilon > 0) \quad \lim_{n\to\infty} \mu(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| \geq \epsilon\}) = 0.$$

Problem 4

(a) From the above definitions, convergence in measure has the limit outside, so it only needs that the points where X_n is far from X to be few. This means that I can choose some points wherever to be far away, and it would still be convergent in measure. However, convergence a.e. has the limit inside, so having bursts of points far away would destroy the convergence of X_n to X, because we can always choose the bursts so that infinitely often X_n would again be far from X.

For an explicit construction, consider $X_1 = 1_{[0,1]}, X_2 = 1_{[0,\frac{1}{2}]}, X_3 = 1_{[\frac{1}{2},1]}, X_4 = 1_{[0,\frac{1}{3}]}, X_5 = 1_{[\frac{1}{3},\frac{2}{3}]}, \ldots$ and so on. These X_n clearly converge in measure to 0 because $\frac{1}{n}$ eventually gets below any ϵ, but X_n to not converge a.e. to 0 because for every $\omega \in [0,1]$, $X_n = 1$ infinitely often.

(b) An example where infinite measure causes problems: consider $X_n = \omega^{1/n}$ on $[1, \infty)$. The X_n converge a.e. to 1 (for every ω, it's possible to find N s.t. when $n \geq N$, $|\omega^{1/n} - 1| < \epsilon$), but there will always be a set of infinite measure that where $|X_n(\omega) - X(\omega)| \geq \epsilon$, and hence $\lim_{n\to\infty} \mu(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| \geq \epsilon\}) = \infty \neq 0$.

Problem 5

One direction: for any $\epsilon > 0$, consider the sets $A = \{\omega \in \Omega : |X_m - X| < \frac{\epsilon}{2}\}$, $B = \{\omega \in \Omega : |X - X_n| < \frac{\epsilon}{2}\}$, and $C = \{\omega \in \Omega : |X_m - X_n| < \epsilon\}$. We are given that the X_i converge in measure to X, which means that $\mu(A^{\complement}) \to 0$ and $\mu(B^{\complement}) \to 0$. Now for all $\omega \in A \cap B$, we know that $|X_m - X_n| \leq |X_m - X| + |X - X_n| < \epsilon$ by the triangle inequality. Thus, $A \cap B \subseteq C \implies (A \cap B)^{\complement} \supseteq C^{\complement}$. But $\mu(C^{\complement}) \leq \mu((A \cap B)^{\complement}) = \mu(A^{\complement} \cup B^{\complement}) \leq \mu(A^{\complement}) + \mu(B^{\complement}) = 0$. Hence, the X_i converge mutually to X.

521 Homework 2

Daniel Rui - 10/9/19

Problem 1

Before we begin (the proof of the actual question is on the next page), I would like to give some of the basic definitions as an aid to my current learning and future referencing:

$\liminf_{n\to\infty} A_n$ is defined to be $\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k$. In English, this is interpreted as: all $\omega \in \Omega$ that are in all the A_k for $k \geq N$ for some $N \in \mathbb{N}$, or: all $\omega \in \Omega$ that are in all *but* finitely many of the A_k.
$\liminf_{n\to\infty} x_n$ for a sequence of numbers is defined to be $\lim_{n\to\infty}\left(\inf_{k\geq n} x_k\right)$

$\limsup_{n\to\infty} A_n$ is defined to be $\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k$. In English, this is interpreted as: all $\omega \in \Omega$ that are in infinitely many A_k, or: all $\omega \in \Omega$ such that $\omega \in A_k$ infinitely often (abbr: i.o.).
$\limsup_{n\to\infty} x_n$ for a sequence of numbers is defined to be $\lim_{n\to\infty}\left(\sup_{k\geq n} x_k\right)$

For the motivation/intuition behind these particular definitions, see this MSE post and in particular Hans Lundmark's answer (quoted below with addendums by me):

For an increasing sequence of sets $L_1 \subseteq L_2 \subseteq \ldots$, it's intuitive that the limit $L_n \to L$ should be defined as the union of all sets in the sequence, $L = \bigcup_{n=1}^{\infty} L_n$. Similarly, for a decreasing sequence $U_1 \supseteq U_2 \supseteq U_3 \supseteq \cdots$, it's natural to define the limit as the intersection: $U_n \to U$ as $n \to \infty$, where $U = \bigcap_{n=1}^{\infty} U_n$.
Now, for an *arbitrary* sequence of sets $\{A_1, A_2, A_3, \ldots\}$, we can squeeze it between an increasing sequence $\{L_n\}$ (a "lower bound") and a decreasing sequence $\{U_n\}$ (an "upper bound"), like this:

$$\begin{array}{rrcccrr}
L_1 = & A_1 \cap A_2 \cap A_3 \cap \cdots & \subseteq & A_1 & \subseteq & U_1 = & A_1 \cup A_2 \cup A_3 \cup \cdots \\
L_2 = & A_2 \cap A_3 \cap \cdots & \subseteq & A_2 & \subseteq & U_2 = & A_2 \cup A_3 \cup \cdots \\
L_3 = & A_3 \cap \cdots & \subseteq & A_3 & \subseteq & U_3 = & A_3 \cup \cdots
\end{array}$$

and so on. Moreover, $\{L_n\}$ is the largest increasing sequence s.t. $L_n \subseteq A_n$ for all n, and $\{U_n\}$ is the smallest decreasing sequence s.t. $A_n \subseteq U_n$ for all n, so it makes sense to define

$$\liminf_{n\to\infty} A_n = \lim_{n\to\infty} L_n, \qquad \limsup_{n\to\infty} A_n = \lim_{n\to\infty} U_n.$$

[Comment:] "largest increasing sequence" $\implies$ if $\{L'_n\}$ is an increasing sequence such that $L_n \subseteq L'_n \subseteq A_n$ for all n, then $L'_n = L_n$ for all n. [...] Suppose that $\{L_n\}$ fulfills the assumptions, but $L_m \subset L'_m$ (strict inclusion) for some m, i.e. L'_m contains some x s.t. $x \notin L_m = A_m \cap A_{m+1} \cap \ldots$. This means that for all $k \geq m$ we have $x \notin A_k$. Then, since the sequence $\{L'_n\}$ is increasing, that element x has to belong to L'_k as well, but then $L'_k \subseteq A_k$ fails; contradiction.

Furthermore, we know that for increasing D_n and decreasing E_n (this is proven literally everywhere, but I'm most familiar with the explanation in Axler's book Chapter 2C):

$$\mu\left(\lim_{n\to\infty} D_n\right) = \mu\left(\bigcup_{n=1}^{\infty} D_n\right) = \lim_{n\to\infty}\mu(D_n), \qquad \mu\left(\lim_{n\to\infty} E_n\right) = \mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty}\mu(E_n)$$

521 Homework 1

Daniel Rui - 10/2/19

Problem 1

(a) We want to show that if $\{\mathcal{A}_n\}$ is an increasing sequence of algebras, then $\bigcup_{n=1}^{\infty} \mathcal{A}_n$ is an algebra. Thus we need to see if that countable union satisfies the three properties of an algebra:

- $\varnothing \in \bigcup_{n=1}^{\infty} \mathcal{A}_n$ because $\varnothing$ is an element of every $\mathcal{A}_n$
- if $A \in \bigcup_{n=1}^{\infty} \mathcal{A}_n$ then $\exists$ be some $i \in \mathbb{N}$ s.t. $A \in \mathcal{A}_i$. Hence $\Omega \setminus A \in \mathcal{A}_i \implies \Omega \setminus A \in \bigcup_{n=1}^{\infty} \mathcal{A}_n$
- if $A, B \in \bigcup_{n=1}^{\infty} \mathcal{A}_n$, then again $\exists i, j \in \mathbb{N}$ s.t. $A \in \mathcal{A}_i$ and $B \in \mathcal{A}_j$. Take $k = \max\{i, j\}$; then because $A_n \subset A_{n+1}$ for all $n \geq 1$ (by definition of increasing sequence), $A, B \in \mathcal{A}_k \implies A \cup B \in \mathcal{A}_k \implies A \cup B \in \bigcup_{n=1}^{\infty} \mathcal{A}_n$

(b) The above result does not necessarily hold true for σ-algebras $\{\mathcal{A}_n\}$ — in essence, the property that σ-algebras are closed under countable unions generates more than what the countable union of increasing σ-algebras can hold:

Let's denote $\mathcal{A} = \bigcup_{n=1}^{\infty} \mathcal{A}_n$. Take $\Omega = \mathbb{N}$ and $\mathcal{A}_n$ to be the σ-algebra on Ω generated by the power set of the natural numbers $\{1, 2, \dots, n\}$. It is clear that $\mathcal{A}_n \subset \mathcal{A}_{n+1}$ for all $n \geq 1$ because $\mathscr{P}(\{1, \dots, n\}) \subset \mathscr{P}(\{1, \dots, n+1\})$, and also that every even natural number $\{2k\}$ is an element of $\mathcal{A} = \bigcup_{n=1}^{\infty} \mathcal{A}_n$ (because $\{2k\} \in \mathcal{A}_{2k}$). Now if $\mathcal{A}$ *were to be* a σ-algebra, then $\bigcup_{k=1}^{\infty} \{2k\} = 2\mathbb{N}$ would be in $\mathcal{A}$. However, $2\mathbb{N}$ is not an element of any of the $\mathcal{A}_n$ (it may be a subset of some element of $\mathcal{A}_n$, but it isn't *exactly* equal to any one of the elements of $\mathcal{A}_n$), and hence it can not be an element of the countable union of such $\mathcal{A}_n$, namely $\mathcal{A}$.

Problem 2

1. If the intersection (denoted $\bigcap_{\mathcal{A}}$) of all possible algebras generated by a collection $\mathcal{C}$ of subsets of Ω is an algebra, then it is the minimal algebra generated by $\mathcal{C}$. We prove this by contradiction: suppose that there was a smaller algebra than $\bigcap_{\mathcal{A}}$, denoted $\mathcal{A}'$. That would mean that there exists some set in $\bigcap_{\mathcal{A}}$ not contained in $\mathcal{A}'$. But that's impossible, because to be in $\bigcap_{\mathcal{A}}$, a set must be in ALL possible algebras generated by $\mathcal{C}$, including of course $\mathcal{A}'$. Now we prove that the intersection is an algebra:

   - $\varnothing \in \bigcap_{\mathcal{A}}$ because $\varnothing$ is an element of every algebra $\mathcal{A}$.
   - Denote $\{\mathcal{A}\}$ as the set of all possible algebras generated by $\mathcal{C}$. Then, $A \in \bigcap_{\mathcal{A}} \implies A \in \mathcal{A}$ for all $\mathcal{A} \in \{\mathcal{A}\} \implies \Omega \setminus A \in \mathcal{A}$ for all $\mathcal{A} \in \{\mathcal{A}\} \implies \Omega \setminus A \in \bigcap_{\mathcal{A}}$.
   - $A_1, A_2 \in \bigcap_{\mathcal{A}} \implies A_1, A_2 \in \mathcal{A}$ for all $\mathcal{A} \in \{\mathcal{A}\} \implies A_1 \cup A_2 \in \mathcal{A}$ for all $\mathcal{A} \in \{\mathcal{A}\} \implies A_1 \cup A_2 \in \bigcap_{\mathcal{A}}$.

2. Likewise, if the intersection $\bigcap_{\mathcal{A}}$ of all possible σ-algebras generated a collection $\mathcal{C}$ of subsets of Ω is a σ-algebra, then it is the minimal σ-algebra generated by $\mathcal{C}$, with the proof exactly the same

as above. As before, we just need to prove that $\bigcap_{\mathcal{A}}$ is actually a σ-algebra (using all the same notation but just referring to σ-algebras instead of algebras):

- $\varnothing \in \bigcap_{\mathcal{A}}$ because $\varnothing$ is an element of every σ-algebra $\mathcal{A}$.
- $A \in \bigcap_{\mathcal{A}} \implies A \in \mathcal{A}$ for all $\mathcal{A} \in \{\mathcal{A}\} \implies \Omega \setminus A \in \mathcal{A}$ for all $\mathcal{A} \in \{\mathcal{A}\} \implies \Omega \setminus A \in \bigcap_{\mathcal{A}}$.
- $A_1, A_2, \ldots \in \bigcap_{\mathcal{A}} \implies A_1, A_2, \ldots \in \mathcal{A}$ for all $\mathcal{A} \in \{\mathcal{A}\} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ for all $\mathcal{A} \in \{\mathcal{A}\} \implies \bigcup_{i=1}^{\infty} \mathcal{A}_i \in \bigcap_{\mathcal{A}}$.

3. And finally, if the intersection $\bigcap_{\mathcal{M}}$ of all possible monotone-classes generated by a collection $\mathcal{C}$ of subsets of Ω is a monotone-class, then it is the minimal monotone-class generated by $\mathcal{C}$, with the proof exactly the same as the previous cases. And like before, we just need to prove that $\bigcap_{\mathcal{M}}$ is actually a monotone-class (using all the same notation but just referring to monotone-classes and $\mathcal{M}$'s instead of (σ-)algebras and $\mathcal{A}$'s):

   - $M_1, M_2, \ldots \in \bigcap_{\mathcal{M}}$ and $M_1 \subset M_2 \subset M_3 \subset \ldots \implies M_1, M_2, \ldots \in \mathcal{M}$ for all $\mathcal{M} \in \{\mathcal{M}\} \implies \bigcup_{i=1}^{\infty} M_i \in \mathcal{M}$ for all $\mathcal{M} \in \{\mathcal{M}\} \implies \bigcup_{i=1}^{\infty} \mathcal{M}_i \in \bigcap_{\mathcal{M}}$.
   - $M_1, M_2, \ldots \in \bigcap_{\mathcal{M}}$ and $M_1 \supset M_2 \supset M_3 \supset \ldots \implies M_1, M_2, \ldots \in \mathcal{M}$ for all $\mathcal{M} \in \{\mathcal{M}\} \implies \bigcap_{i=1}^{\infty} M_i \in \mathcal{M}$ for all $\mathcal{M} \in \{\mathcal{M}\} \implies \bigcap_{i=1}^{\infty} \mathcal{M}_i \in \bigcap_{\mathcal{M}}$.

Problem 3

Because $\sigma[\mathcal{C}]$ for any collection $\mathcal{C}$ is defined as the minimal sigma-algebra (i.e. $\sigma[\mathcal{C}]$ is the intersection of all sigma-algebras containing $\mathcal{C} \implies$ if something is in $\sigma[\mathcal{C}]$ then it must be in all other σ-algebras containing $\mathcal{C}$), we see that $\mathcal{C}_1 \subset \sigma[\mathcal{C}_2] \implies \sigma[\mathcal{C}_1] \subseteq \sigma[\mathcal{C}_2]$ and symmetrically $\mathcal{C}_2 \subset \sigma[\mathcal{C}_1] \implies \sigma[\mathcal{C}_2] \subseteq \sigma[\mathcal{C}_1]$, so $\sigma[\mathcal{C}_1] = \sigma[\mathcal{C}_2]$.

Problem 5

The binomial distribution is

$$\text{Binom}(n, p_n) : P(X_n = k) = \binom{n}{k} (p_n)^k (1 - p_n)^{n-k}$$

We are given that $\lim\limits_{n\to\infty} np_n = \lambda > 0$. I'll write down the limit properties I'll be using, just for sake of completeness. If $\lim\limits_{x\to c} f(x)$ and $\lim\limits_{x\to c} g(x)$ exist, then:

- The limit of the sum, difference, product, quotient is the sum, difference, product, quotient of the limit ($\lim_{x\to c} g(x)$ can't equal 0 in quotient rule). I'll mainly be using the product rule.
- $\lim\limits_{x\to c}[f(g(x))] = f[\lim\limits_{x\to c} g(x)]$ if $f(x)$ is continuous at $\lim\limits_{x\to c} g(x)$. Below I use the fact that $\exp(x)$ and x^{-k} for any $k \in \mathbb{N}$ is continuous at whatever point is being considered at the time.

Below, I also use the Taylor expansion of $\ln(1 - x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \cdots$.

A final note: we fix some finite constant k before we take the limit of n.

$$\lim_{n\to\infty}\left[\binom{n}{k}(p_n)^k(1-p_n)^{n-k}\right] = \lim_{n\to\infty}\left[\binom{n}{k}\left(\frac{p_n}{1-p_n}\right)^k(1-p_n)^n\right]$$

$$= \lim_{n\to\infty}\left[\binom{n}{k}\left(\frac{p_n}{1-p_n}\right)^k\right]\cdot\lim_{n\to\infty}\left[\left(1-\frac{np_n}{n}\right)^n\right]$$

$$= \lim_{n\to\infty}\left[\binom{n}{k}\left(\frac{1}{p_n}-1\right)^{-k}\right]\cdot\lim_{n\to\infty}\left[\exp\left(\ln\left[\left(1-\frac{np_n}{n}\right)^n\right]\right)\right]$$

$$= \lim_{n\to\infty}\left[\binom{n}{k}\left(\frac{n}{np_n}-\frac{n}{n}\right)^{-k}\right]\cdot\lim_{n\to\infty}\left[\exp\left(n\ln\left[\left(1-\frac{np_n}{n}\right)\right]\right)\right]$$

$$= \lim_{n\to\infty}\left[\binom{n}{k}\cdot n^{-k}\left(\frac{1}{np_n}-\frac{1}{n}\right)^{-k}\right]\cdot\exp\left(\lim_{n\to\infty}\left[n\ln\left(1-\frac{np_n}{n}\right)\right]\right)$$

$$= \lim_{n\to\infty}\left[\frac{\binom{n}{k}}{n^k}\right]\cdot\lim_{n\to\infty}\left[\left(\frac{1}{np_n}-\frac{1}{n}\right)^{-k}\right]\cdot\exp\left(\lim_{n\to\infty}\left[n\left(-\frac{np_n}{n}-\frac{(np_n)^2}{2n^2}-\frac{(np_n)^3}{3n^3}-\cdots\right)\right]\right)$$

$$= \lim_{n\to\infty}\left[\frac{\binom{n}{k}}{n^k}\right]\cdot\left(\lim_{n\to\infty}\left[\frac{1}{np_n}-\frac{1}{n}\right]\right)^{-k}\cdot\exp\left(\lim_{n\to\infty}\left[-np_n-\frac{(np_n)^2}{2n}-\frac{(np_n)^3}{3n^2}-\cdots\right]\right)$$

$$= \lim_{n\to\infty}\left[\frac{\binom{n}{k}}{n^k}\right]\cdot\left(\frac{1}{\lambda}\right)^{-k}\cdot e^{-\lambda} = \frac{1}{k!}\lim_{n\to\infty}\left[\frac{n\cdot(n-1)\cdots(n-k+1)}{n^k}\right]\cdot\lambda^k\cdot e^{-\lambda}$$

$$=\frac{1}{k!}\lim_{n\to\infty}\left[\left(\frac{n}{n}\right)\cdot\left(\frac{n-1}{n}\right)\cdots\left(\frac{n-k+1}{n}\right)\right]\cdot\lambda^k\cdot e^{-\lambda}$$

$$=\frac{1}{k!}\left[\lim_{n\to\infty}\left(\frac{n}{n}\right)\right]\cdot\left[\lim_{n\to\infty}\left(\frac{n-1}{n}\right)\right]\cdots\left[\lim_{n\to\infty}\left(\frac{n-k+1}{n}\right)\right]\cdot\lambda^k\cdot e^{-\lambda} = \frac{\lambda^k e^{-\lambda}}{k!}$$

and finally, we are done.