

Probability, homework 5, solutions.

Exercise 1. Prove that if a sequence of real random variables (X_n) converge in distribution to X , and (Y_n) converges in distribution to a constant c , then $X_n + Y_n$ converges in distribution to $X + c$.

Solution. Let's do it with ε 's and δ 's. Let f be a continuous bounded function, eventually chosen of type $e^{i\lambda \cdot}$. Then

$$\mathbb{E} f(X_n + Y_n) = \mathbb{E} f(X_n + Y_n) \mathbf{1}_{|Y_n - c| < \varepsilon} + \mathbb{E} f(X_n + Y_n) \mathbf{1}_{|Y_n - c| \geq \varepsilon}.$$

The second expectation on the RHS above is bounded by $\|f\|_\infty \mathbb{P}(|Y_n - c| \geq \varepsilon) \leq \|f\|_\infty \mathbb{E} h(Y_n)$ where h is 0 on $[c - \varepsilon/2, c + \varepsilon/2]$, 1 in $[c - \varepsilon, c + \varepsilon]^c$ and in $(0, 1)$ otherwise. As $n \rightarrow \infty$ we have $\mathbb{E} h(Y_n) \rightarrow \mathbb{E} h(c) = 0$ so that we just need to prove

$$|\mathbb{E} f(X_n + Y_n) \mathbf{1}_{|Y_n - c| < \varepsilon} - \mathbb{E} f(X_n + c) \mathbf{1}_{|Y_n - c| < \varepsilon}| \leq C\varepsilon \quad (0.1)$$

for some $C > 0$ which may depend on f but not ε . As we can choose $f(x) = e^{i\lambda x}$ for some fixed $\lambda \in \mathbb{R}$, f is Lipschitz with constant λ^{-1} so that

$$|\mathbb{E} f(X_n + Y_n) \mathbf{1}_{|Y_n - c| < \varepsilon} - \mathbb{E} f(X_n + c) \mathbf{1}_{|Y_n - c| < \varepsilon}| \leq \frac{\varepsilon}{\lambda}. \quad (0.2)$$

Moreover, by the same reasoning as 6 lines above,

$$|\mathbb{E} f(X_n + c) - \mathbb{E} f(X_n + c) \mathbf{1}_{|Y_n - c| < \varepsilon}| \leq \|f\|_\infty \mathbb{P}(|Y_n - c| \geq \varepsilon) \rightarrow 0. \quad (0.3)$$

We also have, by convergence in distribution of X_n ,

$$\mathbb{E} f(X_n + c) \rightarrow \mathbb{E} f(X + c). \quad (0.4)$$

Equations (0.2), (0.3), (0.4) allow to conclude that (0.1) holds.

Exercise 2. Assume that (X, Y) has joint density

$$ce^{-(1+x^2)(1+y^2)},$$

where c is properly chosen. Are X and Y Gaussian random variables? Is (X, Y) a Gaussian vector?

Solution. The answer is that X and Y are not Gaussian random variables. Below is why. Note that for any $c > 0$ we have

$$\int_{\mathbb{R}} e^{-cy^2} dy = \sqrt{\pi/c},$$

as seen from a change of variables from the density of a standard Gaussian.

The density of X at point x is therefore

$$\int ce^{-(1+x^2)(1+y^2)} dy = ce^{-(1+x^2)} \int e^{-(1+x^2)y^2} dy = c\sqrt{\pi/(1+x^2)}e^{-x^2}.$$

Note that this is not the density of a Gaussian random variable.

Moreover, the couple (X, Y) is not a Gaussian vector because if it was, its log-density would be a quadratic form in (x, y) , but it is not (it is a quartic form). Of course it is not a Gaussian vector also because otherwise its coordinates would be Gaussian.

Exercise 3. Let $\epsilon > 0$ and X be uniformly distributed on $[0, 1]$. Prove that, almost surely (i.e. the following event has probability 1), there exists only a finite number of rationals $\frac{p}{q}$, with $p \wedge q = 1$, such that

$$\left| X - \frac{p}{q} \right| < \frac{1}{q^{2+\epsilon}}.$$

Solution. Let $A_{p,q} = \left\{ \left| X - \frac{p}{q} \right| < \frac{1}{q^{2+\epsilon}} \right\}$. Then $\mathbb{P}(A_{p,q}) < 2/q^{2+\epsilon}$ so that

$$\sum_{q \geq 1, p \leq q, p \wedge q = 1} \mathbb{P}(A_{p,q}) \leq \sum_{q \geq 1, p \leq q} \mathbb{P}(A_{p,q}) = \sum_{q \geq 1} \frac{q}{q^{2+\epsilon}} < \infty.$$

This implies (e.g. by Borel-Cantelli) that $\mathbb{P}(A_{p,q} \text{ with some } p \wedge q = 1 \text{ i.o.}) = 0$.

Exercise 4. You toss a coin repeatedly and independently. The probability to get a head is p , a tail is $1-p$. Let A_k be the following event: k or more consecutive heads occur amongst the tosses numbered $2^k, \dots, 2^{k+1} - 1$. Prove that $\mathbb{P}(A_k \text{ i.o.}) = 1$ if $p \geq 1/2$, 0 otherwise.

Here, i.o. stands for “infinitely often”, and A_k i.o. is the event $\cap_{n \geq 1} \cup_{m \geq n} A_m$.

Solution. Let X_i 's be i.i.d Bernoulli's with parameter p , write $X_i = 1$ if the i -th coin toss is a head, 0 otherwise.

Assume $p < 1/2$ first. Let $B_k^\ell = \cap_{i=\ell}^{\ell+k-1} \{X_i = 1\}$. Then $A_k \subset \cup_{\ell=2^k}^{2^{k+1}} B_k^\ell$, so

$$\mathbb{P}(A_k) \leq \sum_{\ell=2^k}^{2^{k+1}} \mathbb{P}(B_k^\ell) \leq 2^{k+1} p^k,$$

so that $\sum_k \mathbb{P}(A_k) < \infty$. This implies easily $\mathbb{P}(A_n \text{ i.o.}) = 0$ (we have now seen this is a part of the Borel-Cantelli lemma).

Assume now that $p \geq 1/2$. Let $C_k^\ell = \{\exists i \in [\ell, \ell + k - 1] : X_i = 0\}$. Then $A_k^c \subset \cap_{a \geq 0: 2^k + ak \leq 2^{k+1}} C_k^{2^k + ak}$, and these events are independent, so

$$\mathbb{P}(A_k^c) \leq \mathbb{P}\left(\cap_{a \geq 0: 2^k + ak \leq 2^{k+1}} C_k^{2^k + ak}\right) = (1 - p^k)^{\lfloor 2^k/k \rfloor} = \exp\left(-\frac{(2p)^k}{k}(1 + o(1))\right).$$

We conclude that $\mathbb{P}(A_k) \geq C/k$ for some universal constant C (this is the bound obtained in the worse case $p = 1/2$). Moreover, by convergence of sets,

$$\mathbb{P}(\cap_{n \geq 1} \cup_{m \geq n} A_m) = \lim_{n \rightarrow \infty} \mathbb{P}(\cup_{m \geq n} A_m) = \lim_{n \rightarrow \infty} \left(1 - \prod_{m \geq n} (1 - \mathbb{P}(A_m))\right).$$

The above parenthesis is always 1 due to $\mathbb{P}(A_k) \geq C/k$, which concludes the proof. For R_n , we just need the upper tail, and the approaches mentioned in (7) all work.

Exercise 5. Prove that for any $x > 0$, $\frac{1}{x} = \int_0^\infty e^{-tx} dt$. Deduce the value of $\int_0^\infty \frac{\sin x}{x} dx$.

Solution. The first question is elementary. For the second question, by Fubini we have

$$\int_0^\infty \frac{\sin x}{x} dx = \int_0^\infty \int_0^\infty e^{-tx} \frac{e^{ix} - e^{-ix}}{2i} dx dt = \int_0^\infty \frac{1}{2i} \left(\frac{1}{t-i} - \frac{1}{t+i} \right) dt = \frac{1}{2} \int_{\mathbb{R}} \frac{dt}{t^2+1} = \frac{\pi}{2},$$

where the last equality is known through the density normalization of a Cauchy random variable (i.e. by a residue calculus) or with the arctan antiderivative.

Exercise 6. For any probability measure μ supported on $\mathbb{R}_+$, one defines the Laplace transform as

$$\mathcal{L}_\mu(\lambda) = \int_0^\infty e^{-\lambda x} d\mu(x), \quad \lambda \geq 0.$$

- (1) Prove that $\mathcal{L}_\mu$ is well-defined, continuous on $\mathbb{R}_+$ and $\mathcal{C}^\infty$ on $\mathbb{R}_+^*$.
- (2) Prove that $\mathcal{L}_\mu$ characterizes the probability measure μ supported on $\mathbb{R}_+$.
- (3) Assume that for a sequence $(\mu_n)_{n \geq 1}$ of probability measure supported on $\mathbb{R}_+$, one has $\mathcal{L}_{\mu_n}(\lambda) \rightarrow \ell(\lambda)$ for any $\lambda \geq 0$, and ℓ is right-continuous at 0. Prove that $(\mu_n)_{n \geq 1}$ is tight, and that it converges weakly to a measure μ such that $\ell = \mathcal{L}_\mu$.

Solution.

- (1) It is well defined because $|\int_0^\infty e^{-\lambda x} d\mu(x)| \leq \int_0^\infty |e^{-\lambda x}| d\mu(x) \leq \int_0^\infty d\mu(x) = 1$. Continuity follows from $|e^{-(\lambda+h)x} - e^{-\lambda x}| \leq 2$ and dominated convergence. Similarly, for infinite differentiability, we prove by induction that

$$\mathcal{L}_\mu^{(k)}(\lambda) = \int_0^\infty (-x)^k e^{-\lambda x} d\mu(x)$$

for all $k \geq 0$ and $\lambda > 0$. The case $k = 0$ is obvious. Assuming it is true for k , for $\lambda, h > 0$ we write

$$(\mathcal{L}_\mu^{(k)}(\lambda+h) - \mathcal{L}_\mu^{(k)}(\lambda))/h = \int_0^\infty (-x)^k \frac{e^{-(\lambda+h)x} - e^{-\lambda x}}{h} d\mu(x).$$

For $\alpha > 0$ we have $|e^{-\alpha} - 1| < \alpha$, so that $|\frac{e^{-(\lambda+h)x} - e^{-\lambda x}}{h}| < x e^{-\lambda x}$ for all h and we can apply dominated convergence as $h \rightarrow 0^+$. A similar reasoning allows to treat the case $h \rightarrow 0^-$ and concludes the induction.

- (2) There are many ways to solve this difficult question. Here is a particularly simple solution. Let ν be the pushforward of μ by the function $x \mapsto e^{-x}$. Then the knowledge of $(\mathcal{L}_\mu^{(k)})_{k \geq 0}$ is equivalent to the knowledge of all moments of ν , a measure supported on $[0, 1]$. Such measures are clearly characterized by their moments, e.g. by the Stone-Weierstrass theorem on $[0, 1]$.
- (3) We first prove tightness. Direct calculation gives, for any $\alpha > 0$,

$$\frac{1}{\alpha} \int_0^\alpha (1 - \mathcal{L}_{\mu_n}(\lambda)) d\lambda = \int_0^\infty \frac{\alpha x - 1 + e^{-\alpha x}}{\alpha x} d\mu_n(x),$$

and $\frac{t-1+e^{-t}}{t}$ is positive and increases on $\mathbb{R}_+$. This implies

$$\frac{1}{\alpha} \int_0^\alpha (1 - \mathcal{L}_{\mu_n}(\lambda)) d\lambda \geq \mu_n([0, \alpha^{-1}]^c).$$

On the other hand,

$$|\frac{1}{\alpha} \int_0^\alpha (1 - \mathcal{L}_{\mu_n}(\lambda)) d\lambda| \leq \frac{1}{\alpha} \int_0^\alpha |1 - \ell(\lambda)| d\lambda + \frac{1}{\alpha} \int_0^\alpha |\ell(\lambda) - \mathcal{L}_{\mu_n}(\lambda)| d\lambda$$

Let $\varepsilon > 0$. By continuity of ℓ at 0 there exists $\alpha > 0$ such that the first integral is at most ε . By dominated convergence, for this α the second integral is also smaller than ε for large enough n . We have proved that

$$\mu_n([0, \alpha^{-1}]^c) \leq \varepsilon$$

for large enough n , i.e. tightness.

As $(\mu_n)_{n \geq 1}$ is tight, it has weakly converging subsequences, to a measure μ supported on $\mathbb{R}_+$. As this weak convergence clearly implies convergence of the Laplace transform, this shows that $\ell = \mathcal{L}_\mu$. Consider any converging subsequence, to a measure μ' . Then $\ell = \mathcal{L}_{\mu'}$, so by (2) $\mu = \mu'$, which concludes the proof.