Imperial College London

COURSEWORK 2

IMPERIAL COLLEGE LONDON

DEPARTMENT OF MATHEMATICS

MATH60028
Probability Theory

Author: Pantelis Tassopoulos

Date: March 9 2023

Problems

Question 1

Let $n \in \mathbb{N}$, $(\lambda_k)_{k\leq n}$ be a collection of non-negative that sums to 1 and $(\zeta_\eta)_{k\leq n}$ be a collection of random variables. Let ϕ_k be the characteristic function of ζ_k for $1 \leq k \leq n$. Show that the function

$$\sum_{k=1}^{n} \lambda_k \cdot \phi_k(t)$$

is a characteristic function as required.

Question 2

Let $\phi(t)$ be a characteristic function corresponding to an arbitrary distribution function $F : \mathbb{R} \to [0,1]$. Using the result from Question one, show that the real part of a characteristic function **is itself, a characteristic function** whereas the imaginary part of a characteristic function can **never be a characteristic function**.

Question 3

Consider n independent fair coin tosses X_n distributed. How many tosses do you need to make to be sure with a 95% chance that my estimate is within 0.01 of the actual value.

Question 4

Suppose you and your friend are tossing fair coins. For $n \in \mathbb{N}$, consider the probability of both of you getting the same number of heads and obtain the asymptotics for these probabilities.

Solutions

Question 1

We want to define the discrete random variable η in the following way:

$$\eta = \begin{cases} 1, & \text{with probability} \lambda_1, \\ \cdots & \\ k, & \text{with probability} \lambda_k, \\ \cdots & \\ n, & \text{with probability} \lambda_n, \end{cases} \tag{1}$$

This defines a random since the λ_k are non-negative and sum to 1. Since η as defined in (1) is discrete, its distribution function is

$$F_\eta(x) = \sum_{k=1}^{n} \lambda_k \cdot \mathbb{1}_{(-\infty,k]}(x), \quad x \in \mathbb{R}$$

From lectures, the characteristic functions ϕ_k are associated with distribution functions $F_k, k = 1, \cdots, n$. Having obtained the distribution functions thereof, applying Kolmogorov's Extension theorem (the consistency relations obviously hold by assuming independence) yields mutually independent random variables ζ_k, η on the probability space $(\mathbb{R}^{n+1}, \mathcal{B}(\mathbb{R}^{n+1}), \mathbb{P})$, where $\mathbb{P} = \otimes_{k=1}^{n} \mathbb{P}_k \otimes \mathbb{P}_\eta$. The probability measures $\mathbb{P}_k, \mathbb{P}_\eta$ are the probability measures induced by the corresponding distribution functions F_K, F_η. Now, we compute the characteristic function of ζ_η (which is a random variable as it is a sum of measurable functions on the measurable sets $\{\eta = k\}$):

$$\phi_{\zeta_\eta}(t) = \mathbb{E}[\exp(it\zeta_\eta)] = \int_\Omega \exp(it\zeta_\eta) d\mathbb{P}$$

$$= \sum_{k=1}^{n} \int_{\{\eta=k\}} \exp(it\zeta_\eta) d\mathbb{P} = \sum_{k=1}^{n} \int_{\{\eta=k\}} \exp(it\zeta_k) d\mathbb{P}$$

since

$$\bigsqcup_{k=1}^{n} \{\eta = k\} = \Omega.$$

with $\Omega = \mathbb{R}^{n+1}$. Furthermore,

$$\sum_{k=1}^{n} \int_{\{\eta=k\}} \exp(it\zeta_k) d\mathbb{P} = \sum_{k=1}^{n} \int_\Omega \exp(it\zeta_k) \mathbb{1}_{\{\eta=k\}}(\omega) d\mathbb{P}(\omega)$$

$$\sum_{k=1}^{n} \int_\Omega \exp(it\zeta_k) \mathbb{1}_k(\eta(\omega)) d\mathbb{P}(\omega) = \sum_{k=1}^{n} \mathbb{E}[\exp(it\zeta_k) \cdot \mathbb{1}_k(\eta)]$$

Now, by the independence of ζ_k with η, the expectations factorise and we obtain

$$\phi_{\zeta_\eta}(t) = \sum_{k=1}^{n} \mathbb{E}[\exp(it\zeta_k) \cdot \mathbb{1}_k(\eta)] = \sum_{k=1}^{n} \mathbb{E}[\exp(it\zeta_k)] \cdot \mathbb{E}[\mathbb{1}_k(\eta)]$$

$$\sum_{k=1}^{n} \mathbb{E}[\exp(it\zeta_k)] \mathbb{P}(\eta = k) = \sum_{k=1}^{n} \lambda_k \cdot \phi_k(t)$$

Thus,

$$\sum_{k=1}^{n} \lambda_k \cdot \phi_k(t)$$

is indeed a characteristic function as required.

Question 2

Let $\phi(t)$ be a characteristic function corresponding to an arbitrary distribution function $F : \mathbb{R} \to [0,1]$. This enables us to construct a random variable X on the probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_F)$, where $\mathbb{P}_F$ is the probability measure induced F, $X : \Omega \to \mathbb{R}$ and $X(\omega) = \omega$, which has distribution F. Now, using the result from Question one, the random variable

$$Z = X \cdot \mathbb{1}_{\{\eta=1\}} - X \cdot \mathbb{1}_{\{\eta=2\}}$$

with η as in (1) where $n = 2$ and $\lambda_1 = \lambda_2 = \frac{1}{2}$, has characteristic function

$$\phi_Z(t) = \frac{1}{2}(\phi_X(t) + \phi_{-X}(t)) = \frac{1}{2}\left(\mathbb{E}[e^{itX}] + \mathbb{E}[e^{-itX}]\right) = \mathbb{E}\left[\frac{e^{itX} + e^{-itX}}{2}\right]$$

$$= \mathfrak{Re}\,\mathbb{E}\left[e^{itX}\right] = \mathfrak{Re}\,\phi_X(t)$$

using the linearity of expectation. So, indeed, one has that the real part of a characteristic function **is itself, a characteristic function**. Now, for $\mathfrak{Im}\,\phi_X(t)$, notice that

$$\mathfrak{Im}\,\phi_X(0) = \mathfrak{Im}\,\mathbb{E}[e^{i0X}] = \mathfrak{Im}\,\mathbb{E}[1] = \mathfrak{Im}\,1 = 0$$

since taking expectations is equivalent to integration with respect to a probability measure. This means that the imaginary part of a characteristic function can **never be a characteristic function** on its own as characteristic functions evaluate to one when $t = 0$, whereas by the above, $\mathfrak{Im}\,\phi(0) = 0$.

Question 3

A coin toss can be modelled using a Bernoulli random variable X with success probability p taking values in $\{0,1\}$. That is, $X = 1$ with probability p; this is then interpreted as the coil landing with heads on top. Now, we consider n independent such tosses X_n distributed as X. The sample mean

$$S_n = \frac{1}{n}\sum_{k=1}^{n} X_k$$

has mean $\mathbb{E}[S_n] = \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}[X_k] = \dfrac{1}{n}np = p$ and variance $\mathrm{Var}[S_n] = \frac{1}{n^2}\sum_{k=1}^{n}\mathrm{Var}[X_k] = \dfrac{1}{n^2}np(1-p) = \dfrac{p(1-p)}{n}$ by the linearity of expectation and the independence of the X_k. Chebychev's inequality now yields for all $\epsilon > 0$,

$$\mathbb{P}\left(|S_n - \mathbb{E}[S_n]| \geq \epsilon\right) = \mathbb{P}\left(|S_n - p| \geq \epsilon\right) \leq \frac{\mathrm{Var}[S_n]}{\epsilon^2}$$

$$= \frac{p(1-p)}{n\epsilon^2} \leq \frac{1}{4n\epsilon^2} \tag{2}$$

since $p \in [0,1]$ implies $p(1-p) \leq \frac{1}{4}$. Now, set $\epsilon = 0.01$, and let $n = \left\lceil \frac{500^2}{5} \right\rceil$. Substituting the above in (2) gives

$$\mathbb{P}(|S_n - p| \geq \epsilon) = \mathbb{P}(|S_n - p| \geq 0.01) \leq \frac{1}{4n\epsilon^2}$$

$$= \frac{1}{4\left\lceil \frac{500^2}{5} \right\rceil \cdot 0.01^2} \leq \frac{1}{\frac{4}{5}(500 \cdot 0.01)^2} = \frac{1}{20} = \frac{5}{100}.$$

Thus,

$$\mathbb{P}(|S_n - p| \leq \epsilon) \geq 1 - \frac{5}{100} = \frac{95}{100}$$

This precisely means that it suffices to make $n = \left\lceil \frac{500^2}{5} \right\rceil = 50000$ tosses to be sure with a 95% chance that my estimate is within 0.01 of the actual value.

Question 4

Assuming independence between tosses, the number of heads thrown by my friend and myself can be modelled as $X_n = \sum_{k=1}^{n} A_k$ and $Y_n = \sum_{k=1}^{n} B_k$ respectively, where the A_k and the B_k are iid Bernoulli random variables with success probability $\frac{1}{2}$ taking values in $\{0,1\}$. This means that X_n and Y_n are iid Binomial random variables with success probability $\frac{1}{2}$. Now, the probability that I obtain the same number of heads as my friend is

$$\mathbb{P}(X_n = Y_n) = \mathbb{P}\left(\bigsqcup_{k=1}^{n} \{X_n = Y_n = k\}\right) = \sum_{k=1}^{n} \mathbb{P}(X_n = k, Y_n = k) = \sum_{k=1}^{n} \mathbb{P}(X_n = k)\mathbb{P}(Y_n = k)$$

by independence of X_n and Y_n. Now, since X_n, Y_n have the binomial distribution:

$$\mathbb{P}(X_n = Y_n) = \sum_{k=1}^{n} \mathbb{P}(X_n = k)\mathbb{P}(Y_n = k) = \sum_{k=1}^{n} \left[\binom{n}{k} \frac{1}{2^n}\right]^2$$

$$= \frac{1}{4^n} \sum_{k=1}^{n} \binom{n}{k}^2 = \frac{\binom{2n}{n}}{4^n}$$

using a standard binomial identity. Now, consider the random variables representing the total number of heads

$$S_n = X_n + Y_n = \sum_{k=1}^{n} A_k + B_k \sim \text{Bin}\left(\frac{1}{2}, 2n\right)$$

being the sum of two iid binomial Bin($\frac{1}{2}, n$) rv's. The moment generating function of $\bar{S}_n = S_n - \mathbb{E}[S_n] = S_n - n$, i.e. S_n upon centering is

$$\phi_{\bar{S}_n}(t) = \mathbb{E}[e^{it\bar{S}_n}] = \mathbb{E}[e^{itS_n}]e^{-int} = \mathbb{E}\left[\exp\left(it\sum_{k=1}^{n}(A_k + B_k)\right)\right]e^{-int}$$

$$= \prod_{k=1}^{n}\mathbb{E}[it(A_k + B_k)]e^{-int} = \prod_{k=1}^{n}\mathbb{E}[itA_k]\mathbb{E}[itB_k]e^{-int}$$

$$= \left(\frac{1}{2} + \frac{1}{2}e^{it}\right)^{2n} e^{-int}$$

by the independence of the A_k and the B_k and using the expression for the characteristic function of a Bernoulli random variable. Now, notice that the integral

$$I_n = \int_{-\pi}^{\pi}\phi_{\bar{S}_n}(t)dt = \int_{-\pi}^{\pi}\left(\frac{1}{2} + \frac{1}{2}e^{it}dt\right)^{2n} e^{-int}dt = \frac{1}{4^n}\sum_{k=0}^{2n}\int_{-\pi}^{\pi}\binom{2n}{k}e^{i(k-n)t}dt$$

$$= \frac{1}{4^n}\sum_{k=0}^{2n}\int_{-\pi}^{\pi}\binom{2n}{k}\cos((n-k)t)dt = 2\pi\frac{\binom{2n}{n}}{4^n}$$

Now, to derive the asymptotics of the above probability, consider

$$\sqrt{n}I_n = \sqrt{n}\int_{-\pi}^{\pi}\left(\frac{1}{2} + \frac{1}{2}e^{it}dt\right)^{2n} e^{-int}dt = \sqrt{n}\int_{-\pi}^{\pi}\left(\cos\frac{t}{2}\right)^{2n}dt$$

The substitution $t \to t\sqrt{\frac{n}{2}}$ yields

$$\sqrt{n}I_n = \sqrt{2}\int_{-\pi\sqrt{\frac{n}{2}}}^{\pi\sqrt{\frac{n}{2}}}\left(\cos\frac{t}{2\sqrt{\frac{n}{2}}}\right)^{2n}dt = \sqrt{2}\int_{-\pi\sqrt{\frac{n}{2}}}^{\pi\sqrt{\frac{n}{2}}}\phi_n(t)dt$$

where ϕ_n is the characteristic function of

$$\frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathrm{Var}[S_n]}}$$

since $\sqrt{\mathrm{Var}[S_n]} = \sqrt{\frac{2n}{4}} = \sqrt{\frac{n}{2}}$ and $\mathbb{E}[S_n] = 2n$, that is the standardised total number of heads in $2n$ coin tosses. Now, by the Central Limit Theorem, one has that for all $t \in \mathbb{R}$,

$$\phi_n(t) \to \exp\left(-\frac{t^2}{2}\right), \quad n \to \infty.$$

I claim that on $J_n = [-\pi\sqrt{\frac{n}{2}}, \pi\sqrt{\frac{n}{2}}]$, one has

$$\phi_n(t) \le \exp\left(-\frac{t^2}{4}\right), \quad t \in J_n \tag{3}$$

This will be accomplished in steps. First, note that is suffices to consider positive t, since both functions in (3) are even in t. Now, on $[0, \frac{\pi}{2}]$,

$$\sin(x) \geq \frac{2}{\pi}x.$$

One can show this elementary inequality by showing that the smooth function $\frac{\sin x}{x}$ (setting its value to one at $x = 0$), is non-increasing on $[0, \frac{\pi}{2}]$. Integrating this inequality from 0 to $t \in [0, \frac{\pi}{2}]$ gives

$$0 \leq \cos(t) \leq 1 - \frac{1}{\pi}t^2, \quad t \in \left[0, \frac{\pi}{2}\right].$$

Thus, on J_n

$$\left(\cos \frac{t}{2\sqrt{\frac{\pi}{2}}}\right)^{2n} \leq \left(1 - \frac{t^2}{2\pi n}\right)^{2n}.$$

Finally, to show that

$$\left(1 - \frac{t^2}{2\pi n}\right)^{2n} \leq \exp\left(-\frac{t^2}{4}\right), \quad t \in J_n,$$

we take natural logarithms of both sides (both sides are positive and the natural logarithm is monotonically increasing and on J_n, $\frac{t^2}{2\pi n} \leq \frac{\pi}{4} < 1$) to obtain

$$2n \ln\left(1 - \frac{t^2}{2\pi n}\right) \leq -\frac{t^2}{4} \iff \ln\left(1 - \frac{t^2}{2\pi n}\right) \leq -\frac{t^2}{8n}$$

Now, another elementary inequality for $t \in (-1, 0]$ is

$$\ln(1 + x) \leq x$$

which can be obtained by proving the inequality $1 + x \leq e^x$ for all $x \in \mathbb{R}$ and taking natural logarithms. Since, $\frac{t^2}{2\pi n} \leq \frac{\pi}{4} < 1$ for $t \in J_n$, we have from the above

$$\ln\left(1 - \frac{t^2}{2\pi n}\right) \leq -\frac{t^2}{2\pi n}$$

and finally,

$$\ln\left(1 - \frac{t^2}{2\pi n}\right) \leq -\frac{t^2}{2\pi n} \leq -\frac{t^2}{8n}, \quad t \in J_n$$

since $\pi < 4$, thereby showing (3). Thus, the functions

$$f_n(t) = \mathbb{1}_{J_n}(t)\phi_n(t) \leq \exp\left(-\frac{t^2}{4}\right), \quad t \in \mathbb{R}$$

Furthermore, as $J_n \uparrow \mathbb{R}$ and $\phi_n(t) \to \exp\left(-\frac{t^2}{2}\right)$ as $n \to \infty$, the f_n converge point-wise to $\exp\left(-\frac{t^2}{2}\right)$ and are majorised by $\exp\left(-\frac{t^2}{4}\right)$, both integrable functions on $\mathbb{R}$. Thus, by Lebesgue's dominated convergence theorem,

$$\sqrt{n}I_n = 2\pi\frac{\binom{2n}{n}}{4^n}\sqrt{n} = \sqrt{2}\int_{\mathbb{R}} f_n(t)dt \to \sqrt{2}\int_{\mathbb{R}} \exp\left(-\frac{t^2}{2}\right)dt = 2\sqrt{\pi}, \quad n \to \infty$$

Thus, we obtain the following asymptotics for the probability of getting the same number of heads in n fair coin tosses between two people:

$$\frac{\sqrt{n}}{4^n}\binom{2n}{n} \to \frac{1}{\sqrt{\pi}}$$

or equally

$$\frac{1}{4^n}\binom{2n}{n} \sim \frac{1}{\sqrt{n}} \implies \frac{1}{4^n}\binom{2n}{n} = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right), \quad n \to \infty.$$

Note that the exact same limiting behaviour can be obtained through the use of Stirling's approximation on the probability

$$\frac{1}{4^n}\binom{2n}{n},$$

the advantage of the previous method is that it is more explicit.