A8: Probability

Sheet 1 — MT21

Chapters 1 and 2

1. A company sells lottery scratch-cards for £1 each. 1% of cards win the grand prize of £50, a further 20% win a small prize of £2, and the rest win no prize at all. Estimate how many cards the company needs to sell to be 99% sure of making an overall profit. [$\Phi(2.3263) = 0.99$]

2. A list consists of 1000 non-negative numbers. The sum of the entries is 9000 and the sum of the squares of the entries of 91000. Let X represent an entry picked at random from the list. Find the mean of X, the mean of X^2, and the variance of X. Using Markov's inequality, show that the number of entries in the list greater than or equal to 50 is at most 180. What is the corresponding bound from applying Markov's inequality to the random variable X^2? What is the corresponding bound using Chebyshev's inequality?

3. For $n \geq 1$, let Y_n be uniform on $\{1, 2, \ldots, n\}$ (i.e. taking each value with probability $1/n$). Draw the distribution function of Y_n/n. Show that the sequence Y_n/n converges in distribution as $n \to \infty$. What is the limit?

4. Let $X_i, i \geq 1$, be i.i.d. uniform on $[0, 1]$. Let $M_n = \max\{X_1, \ldots, X_n\}$.
   (a) Show that $M_n \to 1$ in probability as $n \to \infty$.
   (b) Show that $n(1 - M_n)$ converges in distribution as $n \to \infty$. What is the limit?

5. (a) What is the distribution of the sum of n independent Poisson random variables each of mean 1? Use the central limit theorem to deduce that

$$e^{-n}\left(1 + n + \frac{n^2}{2!} + \cdots + \frac{n^n}{n!}\right) \to \frac{1}{2} \text{ as } n \to \infty.$$

   (b) Let $p \in (0, 1)$. What is the distribution of the sum of n independent Bernoulli random variables with parameter p? Let $0 \leq a < b \leq 1$. Use appropriate limit theorems to determine how the value of

$$\lim_{n\to\infty} \sum_{r\in\mathbb{N}\colon\, an\leq r<bn} \binom{n}{r} p^r (1-p)^{n-r}$$

   depends on a and b.

   [*You may remember answering some parts of this question in Prelims, but should now make sure you cover particularly the cases not included in Prelims.*]

Matthias Winkel: winkel@stats.ox.ac.uk or winkel@maths.ox.ac.uk

6. (a) Let X_n, $n \geq 1$, be a sequence of random variables defined on the same probability space. Show that if $X_n \to c$ in distribution, where c is a constant, then also $X_n \to c$ in probability.

   (b) Show that if $\mathbb{E}\,|X_n - X| \to 0$ as $n \to \infty$, then $X_n \to X$ in probability. Is the converse true?

7. A gambler makes a long sequence of bets against a rich friend. The gambler has initial capital C. On each round, a coin is tossed; if the coin comes up tails, he loses 30% of his current capital, but if the coin comes up heads, he instead wins 35% of his current capital.

   (a) Let C_n be the gambler's capital after n rounds. Write C_n as a product $CY_1Y_2 \dots Y_n$ where Y_i are i.i.d. random variables. Find $\mathbb{E}C_n$.

   (b) Find the median of the distribution of C_{10} and compare it to $\mathbb{E}C_{10}$.

   (c) Consider $\log C_n$. What does the law of large numbers tell us about the behaviour of C_n as $n \to \infty$? How is this consistent with the behaviour of $\mathbb{E}C_n$?

8. Let $\mathbb{H}_n$ be the n-dimensional cube $[-1,1]^n$. For fixed $x \in \mathbb{R}$, show that the proportion of the volume of $\mathbb{H}_n$ within distance $(n/3)^{1/2} + x$ of the origin converges as $n \to \infty$, and find the limit.

   [*Hint: Consider a random point whose n coordinates are i.i.d. with Uniform$[-1,1]$ distribution. If $A \subset \mathbb{H}_n$, then vol(A)/vol$(\mathbb{H}_n)$ is the probability that such a point falls in the set A. Let D_n represent the distance of such a point from the origin; apply an appropriate limit theorem to D_n^2.*]

9. Let $Y_1, Y_2, \dots$ be i.i.d. and uniformly distributed on the set $\{1, 2, \dots, n\}$. Define $X^{(n)} = \min\{k \geq 1 \colon Y_k = Y_j \text{ for some } j < k\}$, the first time that we see a repetition in the sequence Y_i. Interpret the case $n = 365$. Prove that $X^{(n)}/\sqrt{n}$ converges in distribution to a limit with distribution function $F(x) = 1 - \exp(-x^2/2)$ for $x > 0$.
   [*Hint: Observe that*

$$\mathbb{P}\left(X^{(n)} > m\right) = \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{m-1}{n}\right).$$

   You may find it useful to use bounds such as $-h - h^2 < \log(1-h) < -h$ for sufficiently small positive h.]

Department of Statistics, University of Oxford
Matthias Winkel: winkel@stats.ox.ac.uk or winkel@maths.ox.ac.uk

Additional problems

10. Let $X_i, i \geq 1$, be i.i.d. random variables with $\mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1) = 1/2$.
    (a) Define $S_n = \sum_{i=1}^n X_i 2^{-i}$. What is the distribution of S_n? Show that the sequence S_n converges almost surely as $n \to \infty$ [*Hint: Cauchy sequences converge*]. What is the distribution of the limit S?
    (b) Define $R_n = \sum_{i=1}^n 2X_i 3^{-i}$. Show again that the sequence R_n converges almost surely. Is the limit a discrete random variable? Is it a continuous random variable? [*Hint: Consider its expansion in base 3.*]

11. Consider a random variable $U \colon \Omega \to \mathbb{R}$ that is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and that is uniformly distributed on $[0, 1]$. For $n \geq 1$, let $B_n = 1$ if $\lfloor 2^n U \rfloor$ is an odd integer and $B_n = 0$ otherwise.
    (a) Show that B_n, $n \geq 1$, is a sequence of i.i.d. random variables with $\mathbb{P}(B_n = 0) = \mathbb{P}(B_n = 1) = 1/2$.
    (b) Apply 10.(a) to suitable subsequences (X_i) of (B_n) to construct two independent random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, each distributed like S.
    (c) Adapt your answer to (b) to construct a sequence of i.i.d. random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, each distributed like S.

    [*In Prelims Probability, and also in Part A Probability, the existence of probability spaces for even a single continuous random variable or for sequences of discrete random variables are assumed without proof. In Part A Integration, the Lebesgue σ-algebra and Lebesgue measure on $\mathbb{R}$ are constructed. By restricting them to $\Omega := [0, 1]$, we obtain a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, on which the random variable $U(\omega) = \omega$ is uniformly distributed on $[0, 1]$. The present problem demonstrates that this probability space is, in fact, also suitable for sequences of i.i.d. discrete or continuous random variables. Problem 10 suggests an alternative approach to suitable probability spaces of the form $\Omega = \{0, 1\}^{\mathbb{N}} := \{(x_i) \colon x_i \in \{0, 1\}, i \in \mathbb{N}\}$. A subtle point for both is that (assuming the Axiom of Choice), $\mathcal{F}$ cannot be chosen to be the entire power set of Ω and the definition/existence of the probability measure $\mathbb{P} \colon \mathcal{F} \to [0, 1]$ is not entirely straightforward.*]

12. Let A_n be the median of $2n+1$ i.i.d. random variables which are uniform on $[0, 1]$. Find the probability density function of A_n. Deduce a convergence in distribution result for the median (appropriately rescaled) as $n \to \infty$ (feel free to argue informally if you like!).

    [*Hint: consider the probability that A_n lies in a small interval $(x, x+\epsilon)$). How does the density at the point $\left(\frac{1}{2} + \frac{a}{\sqrt{n}}\right)$ behave as $n \to \infty$? (Stirling's formula may be useful).*]

Matthias Winkel: winkel@stats.ox.ac.uk or winkel@maths.ox.ac.uk