Probability 1, Autumn 2015, Problem sheet 11

To be discussed on the week 14 Dec. . . 18 Dec.
Problems marked with "PrCl" are discussed in the large problem class on Thursday 10 Dec. or 17 Dec.
No mandatory HW's as the end of the TB approaches. Nevertheless this is an important sheet.
Solutions will be available on Blackboard on the 19th Dec.

11.1 Let X have moment generating function $M(t)$, and let $\Psi(t) = \ln M(t)$. Show that

$$\Psi(t)|_{t=0} = 0, \qquad \Psi'(t)|_{t=0} = \mathbf{E}(X), \qquad \Psi''(t)|_{t=0} = \mathbf{Var}(X).$$

11.2 Calculate the moment generating function of the Geom(p) distribution by direct computation.

11.3 Let $X \sim$ Geom(p), the number of trials until the first success in a sequence of independent experiments with success probability p. Let I be the indicator of the success of the first trial.

(a) What is the distribution of $(X \mid I = 1)$?

(b) What is the distribution of $(X \mid I = 0)$? (*Mind the memoryless property.*)

(c) By the Tower rule,

$$M(t) = \mathbf{E}e^{tX} = \mathbf{E}\mathbf{E}\big(e^{tX} \mid I\big).$$

Expand the right hand-side using your previous answers, then solve this equation for $M(t)$, thus determining the moment generating function of the Geom(p) distribution.

11.4 (a) Let $X \sim \mathrm{U}(\alpha, \beta)$. Determine its moment generating function $M_X(t)$.

(b) Let Y be the number shown after rolling a fair die. Determine its moment generating function $M_Y(t)$.

(c) Now let $Z \sim \mathrm{U}(0, 1)$ independent of the above Y and, using $M_{Y+Z}(t) = M_Y(t) \cdot M_Z(t)$, conclude that $Y + Z \sim \mathrm{U}(1, 7)$.

PrCl **11.5** An example in the lecture is "An astronomer measures the unknown distance μ of an astronomical object. He performs n i.i.d. measurements each with mean μ and standard deviation 2 lightyears. How large should n be to have ± 0.5 lightyears accuracy with at least 95% probability?". Using the CLT it is shown in the lecture that the astronomer needs 62 measurements. Work out the estimation using Chebyshev's inequality on the sample mean $\bar{X}$. How does your bound compare to 62?

11.6 A fair coin is flipped 60 times, X denotes the number of Heads. Give an upper bound on the probability $\mathbf{P}\{|X - 30| \geq 20\}$ using Chebyshev's inequality.

11.7 A fair coin is flipped 60 times, X denotes the number of Heads. Give an upper bound on the probability $\mathbf{P}\{|X - 30| \geq 20\}$ using a Chernoff bound along the following lines:

(a) Let $Y_t = e^{tX}$, where $0 < t$. Show that $\mathbf{E}Y_t = 2^{-60}(1 + e^t)^{60}$.

(b) Give an upper bound on the probability $\mathbf{P}\{X \geq 50\}$ by applying Markov's inequality on the variable $Y_t \geq 0$.

(c) Find the value of t that makes the above bound the sharpest. (The problem eventually reduces to the minimising problem of the convex function $f(t) = \ln(1 + e^t) - \frac{5}{6}t$.)

(d) Prove that $\mathbf{P}\{|X - 30| \geq 20\} \leq 2 \cdot 3^{60} \cdot 5^{-50} < 10^{-6}$. Compare this with the result of the previous problem.

11.8 On the (simplified version of the) game Roulette, a player bets £1, and looses his bet with probability 19/37, but is given his bet and an extra pound back with probability 18/37. Use the Weak Law of Large Numbers to find the probability that the casino looses money with this game on the (very) long run. Explain your answer in details.

11.9 **Monte Carlo integration.** Let $f : [0, 1] \to \mathbb{R}$ be a square integrable function. Make the notations $I := \int_0^1 f(x)\,dx$, $J := \int_0^1 |f(x)|^2\,dx$. Here is a method to numerically estimate the integral I: Let $U_1, U_2, \ldots$ be i.i.d. $\mathrm{U}(0, 1)$ random variables and $I_n := \big(f(U_1) + f(U_2) + \cdots + f(U_n)\big)/n$ our n^{th} (random) approximation.

(a) Show that $I_n \to I$ in the sense seen at the WLLN (this is called *in probability*).

(b) Use Chebyshev's inequality to estimate the probability $\mathbf{P}\big\{|I_n - I| > \frac{a}{\sqrt{n}}\big\}$ of error ($a > 0$ is fixed, $n \to \infty$).

PrCl **11.10** What is the approximate probability that the sum of 50 independent and identically distributed random variables falls in the interval [0, 30] if the distribution of these variables

(a) is uniform;

(b) has density function $f(x) = 2x$

on the interval [0, 1]?

11.11 Approximate the probability that the sum of 10 000 rolls of a fair die falls between 34 800 and 35 200.

11.12 A die is continually rolled until the total sum of all rolls exceeds 300. Approximate the probability that at least 80 rolls are necessary.

11.13 One has 100 lightbulbs whose lifetimes are independent exponentials with mean 5 hours. If the bulbs are used one at a time, with a failed bulb being replaced immediately by a new one, approximate the probability that there is still a working bulb after 525 hours.

11.14 In the previous problem suppose that it takes a random time, uniformly distributed over (0, 0.5), to replace a failed bulb. Approximate the probability that all bulbs fail by time 550.

11.15 50 real numbers are rounded to integers, and we can assume that the rounding errors are i.i.d. U(−0.5, 0.5) random variables. Estimate the probability that these rounding errors add up to 3 or more (in either direction) when the sum of the 50 numbers is considered.

11.16 The mean mark of a student on an exam is 74, and the standard deviation is 14 marks. 100 students take this exam. Estimate the probability that the average mark on this exam exceeds 75.

11.17 A fair die is rolled 100 times, the outcome if the i^{th} roll is X_i. Estimate

$$\mathbf{P}\Big\{\prod_{i=1}^{100} X_i \le a^{100}\Big\}$$

for any $1 < a < 6$. *HINT: Think logarithm.*

11.18 Repairing a certain type of gadgets requires two independent steps that each take an Exponential amount of time: the first with mean 12 minutes and the second with mean 18 minutes.

(a) Estimate the probability that a repairman can fix 20 of these gadgets within his 8 hours workday.

(b) How many gadgets can he fix with probability at least 95% within 8 hours?