MAST20004 Probability

Tutorial Set 10

1. Let $X \stackrel{d}{=} \mathrm{R}(0,1)$ and $Y \stackrel{d}{=} \mathrm{R}(1,3)$ be independent random variables.
   (a) Compute the moment generating function $M_X(t)$ of X.
   (b) Compute the moment generating function of Y.
   (c) Compute the moment generating function of $Z = X - 2Y + 2$.
   (d) Use the moment generating function $M_X(t)$ to verify that $\mathbb{E}(X) = 1/2$ and $V(X) = 1/12$.

   Solution:

   (a) For $t \neq 0$,

$$\begin{aligned} M_X(t) &= \int_0^1 e^{tx} dx \\ &= \left[\frac{e^{tx}}{t}\right]_0^1 \\ &= \frac{e^t - 1}{t}. \end{aligned}$$

   (b) For $t \neq 0$,

$$\begin{aligned} M_Y(t) &= \int_1^3 \frac{e^{tx}}{2} dx \\ &= \left[\frac{e^{tx}}{2t}\right]_1^3 \\ &= \frac{e^{3t} - e^t}{2t}. \end{aligned}$$

   (c) For $t \neq 0$,

$$\begin{aligned} M_Z(t) &= \mathbb{E}(e^{tZ}) \\ &= \mathbb{E}(e^{t(X-2Y+2)}) \\ &= e^{2t}\mathbb{E}(e^{tX})\mathbb{E}(e^{-2tY}) \\ &= e^{2t} M_X(t) M_Y(-2t) \\ &= \frac{-e^{2t}(e^t - 1)(e^{-6t} - e^{-2t})}{4t^2}. \end{aligned}$$

   (d) For $t \neq 0$,

$$\begin{aligned} M_X'(t) &= \frac{te^t - e^t + 1}{t^2} \\ \mathbb{E}(X) &= \lim_{t \to 0} M_X'(t) \\ &= \lim_{t \to 0} \frac{e^t + te^t - e^t}{2t} \\ &= \frac{1}{2}, \end{aligned}$$

by L'Hopital's Rule.

$$\begin{aligned} M''_X(t) &= \frac{t^3e^t - 2t^2e^t + 2te^t - 2t}{t^4} \\ \mathbb{E}(X^2) &= \lim_{t\to 0} M''_X(t) \\ &= \lim_{t\to 0} \frac{t^3e^t + t^2e^t - 2te^t + 2e^t - 2}{4t^3} \\ &= \lim_{t\to 0} \frac{t^3e^t + 4t^2e^t}{12t^2} \\ &= \frac{1}{3}, \end{aligned}$$

again by L'Hopital's Rule. So $V(X) = 1/3 - 1/4 = 1/12$.

2. Let $Y_\lambda \stackrel{d}{=} \mathrm{Pn}(\lambda)$.
   (a) Write down the moment generating function of Y_λ.
   (b) Compute the moment generating function of $Z_\lambda = (Y_\lambda - \lambda)/\sqrt{\lambda}$.
   (c) Using part (b) show that $Z_\lambda \stackrel{d}{\to} N(0,1)$ as $\lambda \to \infty$.

   Solution:

   (a) For any $t \in \mathbb{R}$,

$$M_{Y_\lambda}(t) = e^{-\lambda(1-e^t)}.$$

   (b) For any $t \in \mathbb{R}$,

$$\begin{aligned} M_{Z_\lambda}(t) &= \mathbb{E}\left(e^{t(Y_\lambda-\lambda)/\sqrt{\lambda}}\right) \\ &= e^{-\sqrt{\lambda}t} M_{Y_\lambda}\left(t/\sqrt{\lambda}\right) \\ &= e^{-\sqrt{\lambda}t} e^{-\lambda\left(1-e^{t/\sqrt{\lambda}}\right)}. \end{aligned}$$

   (c) For any $t \in \mathbb{R}$,

$$\begin{aligned} \log(M_{Z_\lambda}(t)) &= -\sqrt{\lambda}t - \lambda\left(1 - e^{t/\sqrt{\lambda}}\right) \\ &= -\sqrt{\lambda}t + \lambda\left(t/\sqrt{\lambda} + t^2/2\lambda + t^3/6\lambda^{3/2}\dots\right) \\ &= \lambda\left(t^2/2\lambda + t^3/6\lambda^{3/2}\dots\right). \end{aligned}$$

   So $\log(M_{Z_\lambda}(t)) \to t^2/2 \implies M_Z(t) \to e^{t^2/2}$ as $\lambda \to \infty$. Therefore $Z_\lambda \stackrel{d}{\to} N(0,1)$ as $\lambda \to \infty$.

3. Let X and Y be independent random variables, with known moment generating functions $M_X(t)$ and $M_Y(t)$ and, Z be such that $\mathbb{P}(Z=1) = 1 - \mathbb{P}(Z=0) = p \in (0,1)$. Compute the moment generating function of the random variable $S = ZX + (1-Z)Y$. [The distribution of S is called a *mixture* of the distributions of X and Y.]

   Hint: If you don't know how/where to start, think about conditioning (in this case, conditioning on Z looks promising).

   Solution:

   $M_S(t|Z=1) = M_X(t)$ and $M_S(t|Z=0) = M_Y(t)$. Therefore

$$\begin{aligned} M_S(t) &= \mathbb{E}(M_S(t|Z)) \\ &= pM_X(t) + (1-p)M_Y(t). \end{aligned}$$

4. (a) Let $A_n = \frac{S_n}{n}$ denote the sample average of n observations on the discrete distribution with probability mass function

$$p_X(x) = \frac{1}{6}, \quad x = 1,2,3,4,5,6.$$

   Compute an approximate value for $\mathbb{P}(3 < A_n < 4)$ for $n = 12$ and for $n = 24$.

   (b) Let $A_n = \frac{S_n}{n}$ denote the sample average of n observations on a gamma distribution $\gamma(2, 1/4)$. Compute an approximate value for $\mathbb{P}(7 < A_n < 9)$ for $n = 64$ and for $n = 128$.

   Solution:

   (a) $\mathbb{E}(X) = 7/2$ and $V(X) = 35/12$ so $\mathbb{E}(S_n/n) = 7/2$ and $V(S_n/n) = 35/12n$. Therefore, using the normal approximation,

$$\begin{aligned} \mathbb{P}(3 < S_{12}/12 < 4) &\approx \mathbb{P}(-6/\sqrt{35} < Z < 6/\sqrt{35}) \\ &= \mathbb{P}(-1.0142 < Z < 1.0142) \\ &= 0.6895, \end{aligned}$$

   where Z is a standard normal random variable. Also

$$\begin{aligned} \mathbb{P}(3 < S_{24}/24 < 4) &\approx \mathbb{P}(-6\sqrt{2}/\sqrt{35} < Z < 6\sqrt{2}/\sqrt{35}) \\ &= \mathbb{P}(-1.4343 < Z < 1.4343) \\ &= 0.8485. \end{aligned}$$

   (b) $\mathbb{E}(X) = 8$ and $V(X) = 32$ so $\mathbb{E}(S_n/n) = 8$ and $V(S_n/n) = 32/n$. Using the normal approximation,

$$\begin{aligned} \mathbb{P}(7 < S_{64}/64 < 9) &\approx \mathbb{P}(-\sqrt{2} < Z < \sqrt{2}) \\ &= \mathbb{P}(-1.4142 < Z < 1.4142) \\ &= 0.8427, \end{aligned}$$

   where Z is a standard normal random variable. Also

$$\begin{aligned} \mathbb{P}(7 < S_{128}/128 < 8) &\approx \mathbb{P}(-2 < Z < 2) \\ &= 0.9545. \end{aligned}$$

5. The performance level of a machine component after n years of use is well described by a Markov chain with states 1 = good, 2 = fair, and 3 = unsatisfactory. The transition probability matrix is given by

$$\boldsymbol{P} = \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0 & 0.7 & 0.3 \\ 0 & 0 & 1 \end{bmatrix}.$$

(a) Of a large number of such components, which are all good initially, what proportion can be expected to be unsatisfactory after two years? After four years? After eight years?

(b) If initially, only 80% of such components are good and the rest are fair, what proportion can be expected to be fair after three years?

(c) Find the stationary distribution of the Markov chain. Explain how one could check your answer using Matlab.

Solution:

(a) The solution is given by the $(1,3)$ components of $\boldsymbol{P}^2$, $\boldsymbol{P}^4$, and $\boldsymbol{P}^8$, respectively.

$$[1\ 0\ 0]\ \boldsymbol{P}^2 = \begin{bmatrix} 0.36 & 0.39 & 0.25 \\ 0 & 0.49 & 0.51 \\ 0 & 0 & 1.0000 \end{bmatrix} \quad - \quad .36\ 0.3 \quad .25$$

$$[1\ 0\ 0]\ \boldsymbol{P}^4 = \begin{bmatrix} 0.1296 & 0.3315 & 0.5389 \\ 0 & 0.2401 & 0.7599 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{P}^8 = \begin{bmatrix} 0.0168 & 0.1226 & 0.8606 \\ 0 & 0.0576 & 0.9424 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus $\mathbb{P}(X_2 = 3 | X_0 = 1) = 0.25$, $\mathbb{P}(X_4 = 3 | X_0 = 1) = 0.5389$, and $\mathbb{P}(X_8 = 3 | X_0 = 1) = 0.8606$.

(b) We have

$$\begin{bmatrix} 0.8 & 0.2 & 0 \end{bmatrix} \boldsymbol{P}^3 = \begin{bmatrix} 0.8 & 0.2 & 0 \end{bmatrix} \begin{bmatrix} 0.216 & 0.381 & 0.403 \\ 0 & 0.343 & 0.657 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.1728 & 0.3734 & 0.4538 \end{bmatrix}.$$

Thus $\mathbb{P}(X_2 = 3) = 0.3734$.

(c) The stationary distribution is the solution to $\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P}$ and $\pi_1 + \pi_2 + \pi_3 = 1$, which is $\boldsymbol{\pi} = (0, 0, 1)$. You could check this in Matlab by computing $\boldsymbol{P}^n$ for a large value of n. You would expect the answer to be

$$\boldsymbol{P}^n = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

6. Let X be a random variable with pdf $f_X(x) = xe^{-x},\ 0 \leq x < \infty$.
   (a) Compute the cumulant generating function of X.
   (b) Derive the first four cumulants of X.
   (c) Calculate the coefficients of skewness and kurtosis of X and comment on your findings.

   Solution:

   (a) For $t < 1$, $M_X(t) = \int_0^\infty e^{tx} f_X(x)dx \underbrace{=}_{y=(1-t)x} \frac{1}{(1-t)^2} \int_0^\infty ye^{-y}dy = \frac{1}{(1-t)^2}$, hence the cgf of X is

$$K_X(t) = \log M_X(t) = -2\log(1-t) = 2\left(t + \frac{t^2}{2} + \frac{t^3}{3} + \frac{t^4}{4} + \dots\right).$$

   (b) Compare the cgf with the general form of cgf $K_X(t) = \sum_{r=1}^\infty \frac{\kappa_r t^r}{r!}$, we obtain

$$\kappa_1 = 2, \quad \kappa_2 = 2, \quad \kappa_3 = 4, \quad \kappa_4 = 12.$$

   (c) $\sigma = \sqrt{2}$, so $\text{Skew}(X) = \frac{\kappa_3}{(\sqrt{2})^3} = \sqrt{2}$, $\text{Kurt}(X) = \frac{\kappa_4}{(\sqrt{2})^4} = 3$. For the normal, both coefficients of skewness and kurtosis are zero, meaning that the pdf of X is positive skew (right skew) and longer tailed.

7. Extreme values are of central importance in risk management and the following two questions provide the fundamental tool used in the extreme value theory.
   (a) Let $X_1, \dots, X_n$ be independent identically distributed (i. i. d.) exp(1) random variables and define $Z_n = \max(X_1, \dots, X_n)$.
      (i) Find the cumulative distribution of Z_n.
      (ii) Calculate the cumulative distribution of $V_n := Z_n - \ln n$.
      (iii) Derive the limit distribution of V_n as $n \to \infty$.
   (b) Let $Y_1, \dots, Y_n$ be i. i. d. random variables with the pdf

$$f(x) = \begin{cases} \gamma x^{\gamma-1}, & 0 < x < 1, \\ 0, & \text{else}, \end{cases}$$

   where $\gamma > 0$ is a parameter. Define $M_n = n^{1/\gamma} \min(Y_1, \dots, Y_n)$.
      (i) Find the cumulative distribution of M_n.
      (ii) Derive the limit distribution of M_n as $n \to \infty$.

   Solution:

   (a) (i)

$$\begin{aligned} F_{Z_n}(x) &= \mathbb{P}(\max(X_1, \dots, X_n) \leq x) \\ &= \mathbb{P}(X_1 \leq x, \dots, X_n \leq x) = \mathbb{P}(X_1 \leq x)^n \\ &= \begin{cases} (1 - e^{-x})^n, & \text{for } x \geq 0, \\ 0, & \text{for } x < 0. \end{cases} \end{aligned}$$

   (ii) $F_{V_n}(x) = \mathbb{P}(Z_n \leq x + \ln n) = \begin{cases} (1 - e^{-x}/n)^n, & \text{for } x \geq -\ln n, \\ 0, & \text{for } x < -\ln n. \end{cases}$

   (iii) The limit distribution is $G(x) = e^{-e^{-x}}$, for $x \in \mathrm{R}$.

(b) (i) The cdf of Y_1 is $F_{Y_1}(y) = \begin{cases} y^\gamma, & \text{for } y \in [0,1], \\ 0, & \text{for } y < 0, \\ 1, & \text{for } y > 1, \end{cases}$ so

$$\begin{aligned} F_{M_n}(x) &= 1 - \mathbb{P}(M_n > x) = 1 - \mathbb{P}(Y_1 > xn^{-1/\gamma}, \ldots, Y_n > xn^{-1/\gamma}) \\ &= \begin{cases} 1 - (1 - x^\gamma/n)^n, & \text{for } xn^{-1/\gamma} \in [0,1], \\ 0, & \text{for } xn^{-1/\gamma} < 0, \\ 1, & \text{for } xn^{-1/\gamma} > 1. \end{cases} \end{aligned}$$

(ii) The limit distribution is $G(x) = \begin{cases} 0, & \text{for } x < 0, \\ 1 - e^{-x^\gamma}, & \text{for } x \geq 0, \end{cases}$ that is G is Weibull$(1, \gamma)$. This explains why the Weibull distribution is so common in risk management.

MAST20004 Probability

Computer Lab 10

In this lab you

- simulate the total amount T claimed from an insurance company in one day and compare your simulation estimates against the theoretical values of $\mathbb{E}(T)$ and $V(T)$.
- illustrate the Central Limit Theorem and convergence in distribution by simulating sequences of distribution functions converging to the distribution function of a standard normal random variable.
- perhaps have time to revise previous computer labs. I remind you all that there will be at least one question in the exam that will test the theory, as opposed to Matlab coding, that we have dealt with in the Computer Labs. You may also be required to interpret Matlab output.

Exercise A - Simulation of insurance company total claims

Suitably modified, the **incomplete** Matlab m-file **Lab10ExA.m** will simulate the total amount claimed from an insurance company in one day. You will need to add a few lines to the program to generate the required distributions. **Lab10ExA.m** produces estimates for $\mathbb{E}(T)$ and $V(T)$ and also plots the empirical pdf for T.

Let the number of claims in one day be $N \stackrel{d}{=} \text{Pn}(10)$ and $X_1, X_2, \cdots$ be random variables representing claim amounts. We assume that N and $X_1, X_2, \cdots$ are independent, with $X_i \stackrel{d}{=} X$ (for all i) for some claim size X. Then $T = \sum_{i=1}^{N} X_i$ is the sum of a (random) number of random variables and represents the total amount claimed in one day.

1. We start with the assumption that $X_i \stackrel{d}{=} \exp(\lambda)$. Using the appropriate formulae from lectures calculate the theoretical values for $\mathbb{E}(T)$ and $V(T)$.
2. Open **Lab10ExA.m** in the m-file editor and add the code required to generate the claims. Run the program for a couple of different values of λ and compare your theoretical answers with the simulation estimates. Also comment on the shape of the empirical pdf for T.
3. Repeat this exercise for a claim distribution $X \stackrel{d}{=} R(10, 20)$.

Exercise B - Central Limit Theorem - Graphs of distribution functions

Suitably modified, the Matlab m-file **Lab10ExB.m** will generate estimates of the distribution functions F_n for the standardised random variables Z_n as defined on lecture slide 527. Z_n standardises the sample sum $S_n = X_1 + \cdots + X_n$ for a sample of size n. The independent observations $X_1, \ldots, X_n$ in the sample all have the same underlying distribution function F with mean μ and variance σ^2.

The Central Limit Theorem tells us that the distribution functions F_n should tend to the distribution function for a standard normal random variable Z, irrespective of the form of the underlying distribution F.

Lab10ExB.m plots a sequence of distribution functions on any one run - you can specify the selected sequence of n values in the 'Initialisation' section of the program. On each plot the selected distribution functions are compared with the distribution function for a standard normal random variable (plotted as a dotted black line). The program starts with the underlying distribution F set to be a discrete Bernoulli distribution with $p = 0.5$. Don't worry too much about understanding the main plot command in the program.

1. Run the program for $n = [3, 5, 10, 50]$. For F Bernoulli, what is the distribution of S_n? Use your answer to check the distribution functions plotted for $n = 3$ and 5 by using an appropriate probability mass function. Check both the height of some jumps and the position at which they occur.

2. Run the program for some other sequences with larger n values and observe the rate of convergence (for n too large the run time for the program might be quite long).

3. Change the value of p to 0.25 by changing the line in the function 'Finverse'. This changes the mean μ and variance σ^2, so you will need to calculate the new values and type them into the 'Initialisation section'. Repeat the above steps and review the rate of convergence as n increases.

4. To change the underlying distribution you need to change the function 'Finverse'. Comment out the first line and uncomment the line to generate observations on $R(0, 1)$. Again you will need to change the values of μ and σ^2 specified in the 'Initialisation section'. The convergence in the Central Limit Theorem is remarkably quick for the symmetric uniform distribution so you need only consider very small values of n.

5. The program can generate observations on any distribution using the inverse transformation method, by appropriate changes to the function 'Finverse'. A line of code for generating exponentially-distributed X with parameter λ is already included - note that the parameter λ is set in the 'Initialisation' section. Uncomment this line and change μ and σ^2 so that they are calculated from the given value for λ. Then investigate the convergence to normality for a variety of different n and λ values. In particular note the asymmetry in the distribution functions for small values of n caused by the skewness in the exponential distribution.

6. (Optional) Add code to generate observations on some other distributions using the inverse transformation method and investigate the convergence to normality.