ISEG - Lisbon School of Economics and Management
2017/2018

Statistics I
Problem set 6 - Special Random Variables
(version: 9/5/2018)

1. If X has the discrete uniform distribution $f(x) = 1/k$ for $x = 1, 2, ..., k$, show that $E(X) = (k+1)/2$ and $Var(X) = (k^2-1)/12$. (**Hints:** $\sum_{i=1}^{k} i = k\,(k+1)\,/2$ and $\sum_{i=1}^{k} i^2 = k\,(k+1)\,(2k+1)\,/6$).

2. A box contains 10 balls, of which 3 are red, 2 are yellow, and 5 are blue. Five balls are randomly selected with replacement. Calculate the probability that fewer than 2 of the selected balls are red.

3. Prove that if $X_1 \sim B(n_1, p)$ and $X_2 \sim B(n_2, p)$ and X_1 and X_2 are independent random variables, then $X_1 + X_2 \sim B(n_1 + n_2, p)$ (**Hint:** Recall that if $X \sim B(n, p)$, then $M_X(t) = \left[(1-p) + pe^t\right]^n$.)

4. Much is made of the fact that certain mutual funds outperform the market year after year (that is, the return from holding shares in the mutual fund is higher than the return from holding a portfolio such as the S&P 500). For concreteness, consider a 10-year period and let the population be the 4170 mutual funds reported in The Wall Street Journal on January 1, 1995. By saying that performance relative to the market is random, we mean that each fund has a $50-50$ chance of outperforming the market in any year and that performance is independent from year to year.

   (a) If performance relative to the market is truly random, what is the probability that any particular fund outperforms the market in all 10 years?
   (b) Find the probability that at least one fund out of 4170 funds outperforms the market in all 10 years. What do you make of your answer?
   (c) Find the probability that at least five funds outperform the market in all 10 years. (**Hint:** use the law of rare events)

5. Let $X \sim B(n, p)$ and $X^* \sim B(n, 1-p)$, show that $P\,(n - X^* = x) = P\,(X = x)$.

6. Just prior to jury selection for O. J. Simpson's murder trial in 1995, a poll found that about 20% of the adult population believed Simpson was innocent (after much of the physical evidence in the case had been revealed to the public). Ignore the fact that this 20% is an estimate based on a subsample from the population; for illustration, take it as the true percentage of people who thought Simpson was innocent prior to jury selection. Assume that the 12 jurors were selected randomly and independently from the population

   (a) Find the probability that the jury had at least one member who believed in Simpson's innocence prior to jury selection.
   (b) Find the probability that the jury had at least two members who believed in Simpson's innocence.

7. Show that if $X_1 \sim Poisson(\lambda_1)$ and $X_2 \sim Poisson(\lambda_2)$ and if X_1 and X_2 are independent random variables, then $X_1 + X_2 \sim Poisson\,(\lambda_1 + \lambda_2)$ (**Hint:** Recall that if $X \sim Poisson(\lambda)$, then $M_X(t) = e^{\lambda\left(e^t-1\right)}$).

8. Past Experience indicates that an average number of 6 customers per hour stop for petrol at a petrol station. Assuming that the number of customers that stop for petrol at a petrol station is a Poisson random variable:

   (a) What is the probability of 3 customers stopping in any hour?

   (b) What is the probability of 3 customers or less in any hour?

   (c) What is the expected value, and standard deviation of the distribution.

9. The average number of trucks arriving on any one day at a truck depot in a certain city is known to be 12. If we assume a Poisson distribution, what is the probability that on a given day fewer than 9 trucks will arrive at this depot?

10. If 2 percent of the books bound at a certain bindery have defective bindings, use the Poisson approximation to the binomial distribution to determine the probability that 5 of 400 books bound by this bindery will have defective bindings.

11. Records show that the probability is 0.00005 that a car will have a flat tire while crossing a certain bridge. Use the Poisson distribution to approximate the binomial probabilities that, among 10000 cars crossing this bridge,

    (a) exactly two will have a flat tire;

    (b) at most two will have a flat tire.

12. A firm is running 500 vending machines of canned drinks in a town. Every vending machine is defective within one week with probability 1/50. In the case that a machine is defective, the firm has to send out a mechanic. In order to decide whether or not a mechanic should be employed permanently it is of interest to know the probability that the number of defective vending machines X is between 5 and 10 during a week. Determine $P[5 \leq X \leq 10]$ (**Hint:** use the law of rare events).

13. Show that if $X \sim U(0,1)$, then $Y = a + (b-a)X \sim U(a,b)$, $a < b$.

14. In certain experiments, the error made in determining the density of a substance is a random variable having a uniform density with $a = -0.015$ and $b = 0.015$. Find the probabilities that such an error will

    (a) be between -0.002 and 0.003;

    (b) exceed 0.005 in absolute value.

15. (Lack of memory of the exponential random variable) Let $X \sim Exp(\lambda)$, prove that $P(X > x + s | X > x) = P(X > s)$ for any $x \geq 0$ and $s \geq 0$.

16. Let $X_i \sim Exp(\lambda_i)$, $i = 1, 2$, be independent random variables. Prove that $Y = \min\{X_1, X_2\} \sim Exp(\lambda_1 + \lambda_2)$. (**Hint:** Note that $P(Y > y) = P(X_1 > y, X_2 > y)$).

17. Let $X_i \sim Exp(0.5)$, $i = 1, 2$, be independent random variables. Find the expected value of $Z = 2Y_1 + Y_2$ where $Y_1 = \min\{X_1, X_2\}$ and $Y_2 = \max\{X_1, X_2\}$.

18. The mileage (in thousands of kilometers) that car owners get with a certain kind of radial tire is a random variable having an exponential distribution with mean 40. Find the probabilities that one of these tires will last

    (a) at least 20000 km;

    (b) at most 30000 km.

19. Compute the following probabilities:

    (a) If Y is distributed $N(1, 4)$, find $P(Y \leq 3)$.

    (b) If Y is distributed $N(3, 9)$, find $P(Y > 0)$.

    (c) If Y is distributed $N(50, 25)$, find $P(40 \leq Y \leq 52)$.

    (d) If Y is distributed $N(0, 1)$, find $P(|Y| > 1.96)$.

20. If X is a random variable having the standard normal distribution and $Y = X^2$, show that $cov(X, Y) = 0$ even though X and Y are evidently not independent.

21. Suppose that the amount of cosmic radiation to which a person is exposed when flying by jet across Europe is a random variable having a normal distribution with a mean of 4.35 mrem and a standard deviation of 0.59 mrem. What is the probability that a person will be exposed to more than 5.20 mrem of cosmic radiation on such a flight?

22. Prove that if the random variables $X_i, i = 1, 2$ have a normal distribution, $X_i \sim N(\mu_i, \sigma_i^2)$, and are independent and if $Y = aX_1 + bX_2 + c$, then $Y \sim N(\mu_Y, \sigma_Y^2)$, where $\mu_Y = a\mu_1 + b\mu_2 + c$ and $\sigma_Y^2 = a^2\sigma_1^2 + b^2\sigma_2^2$. (**Hint:** Recall that if $X \sim N(\mu, \sigma^2)$, then $M_X(t) = e^{\left(\mu t + 0.5\sigma^2 t^2\right)}$ and note that functions of independent random variables are also independent).

23. In a photographic process, the developing time of prints may be looked upon as a random variable having the normal distribution with $\mu = 15.40$ seconds and $\sigma = 0.48$ second. Find the probabilities that the time it takes to develop one of the prints will be

    (a) at least 16.00 seconds;

    (b) at most 14.20 seconds;

    (c) anywhere from 15.00 to 15.80 seconds.

24. Compute the following probabilities:

    (a) If Y is distributed $\chi^2(4)$ find $P(Y \leq 7.78)$.
    (b) If Y is distributed $\chi^2(10)$ find $P(Y > 18.31)$.
    (c) If Y is $\chi^2(1)$ find $P(Y \leq 3.8416)$.
    (d) If X is distributed $N(0,1)$, find $P(X^2 \leq 3.8416)$.
    (e) Why are the answers to (c) and (d) the same?

25. Show that the square of a standard normal distribution follows a chi-square distribution with 1 degree of freedom (**Hint:** Note that $\Gamma(1/2) = \sqrt{\pi}$).

26. In a certain city, the daily consumption of electric power in millions of kilowatt-hours can be treated as a random variable having a gamma distribution with $a = 3$ and $b = 2$. If the power plant of this city has a daily capacity of 12.59 million kilowatt-hours, what is the probability that this power supply will be inadequate on any given day?

27. Using the moment generating function, show that if $X \sim Gamma(a, b)$ and $Y = 2X/b$, then $Y \sim \chi^2(2a)$.

28. Prove that if X_1 and X_2 are independent random variables with Gamma distribution $X_1 \sim Gamma(a_1, b)$ and $X_2 \sim Gamma(a_2, b)$, then $X_1 + X_2 \sim Gamma(a_1 + a_2, b)$. (**Hint:** Recall that if $X \sim Gamma(a, b)$, then $M_X(t) = (1 - bt)^{-a}$ for $t < 1/b$).

29. The time elapsed since failure until repair (designated as repair time) of a certain type of machines is a random variable with exponential distribution with mean of 2 hours.

    (a) What is the probability that a broken machine has a repair time of 1 hour or less?
    (b) If 10 broken machines were randomly selected, what is the probability of the fastest repair be performed in less than 15 minutes? (assume independence)
    (c) What is the probability that the total repair time of 50 broken machines does not exceed 90 hours? (assume independence)

30. Let X be a Bernoulli random variable with $P(X = 1) = 0.99$, Y be distributed $N(0,1)$ and W be distributed $N(0,100)$. Assume also that X, Y and W are mutually independent random variables and let $S = XY + (1 - X)W$.

    (a) What are the values of $E(Y^2)$ and $E(W^2)$?
    (b) What are the values of $E(Y^3)$ and $E(W^3)$? (**Hint:** What is the skewness for a symmetric distribution?)
    (c) What are the values are $E(Y^4)$ and $E(W^4)$? (**Hint:** use the fact that the kurtosis is 3 for the normal distribution.)

(d) Derive $E(S)$, $E(S^2)$, $E(S^3)$, $E(S^4)$. (**Hint:** Use the law of iterated expectations conditioning on $X = 0$ and $X = 1$)

(e) Derive the skewness and kurtosis for S.

31. A soft-drink vending machine is set so that the amount of drink dispensed is a random variable with a mean of 200 milliliters and a standard deviation of 15 milliliters. Suppose you collect information about 36 vending machines chosen at random. What is the probability that the average amount dispensed by these machines is at least 204 milliliters? (assume independence)

32. Suppose the taxi and takeoff time for commercial jets is a random variable X with a mean of 8.5 minutes and a standard deviation of 2.5 minutes. Assuming independence, what is the probability that for these 36 jets on a given runway total taxi and takeoff time will be

    (a) less than 320 minutes?

    (b) more than 275 minutes?

    (c) between 275 and 320 minutes?

33. Suppose that a book with 300 pages contains on average 1 misprint per page. Assume that the number of misprints per page is a Poisson random variable. What is the probability that there will be at least 100 pages which contain 2 or more misprints? (assume independence)

34. The actual proportion of families in a certain city who own, rather than rent, their home is 0.70. If 84 families in this city are interviewed at random and their responses to the question of whether they own their home are looked upon as values of independent random variables having identical Bernoulli distributions with the parameter $p = 0.70$, with what probability can we assert that the value we obtain for the sample proportion will fall between 0.64 and 0.76? (assume independence)

35. A particular college cannot accommodate more than 1060 students. Assume that each student that applies to the college accepts with probability 0.6. If the college accepts 1700, what is the probability that it will have too many acceptances? (assume independence)

36. Determine the probabilities that the proportion of heads will be anywhere from 0.49 to 0.51 when a balanced coin is flipped

    (a) 100 times;

    (b) 1000 times;

    (c) 10000 times.

37. Suppose that the probability that a person is color blind is 0.005. Estimate the probability that there will be more than one color blind person in 600 people. (**Hint:** Use the law of rare events)

38. Forty seven digits are chosen at random and with replacement from $\{0, 1, 2, ..., 9\}$. Estimate the probability that their average lies between 4 and 5.

39. Suppose that you roll a balanced die 36 times. Let Y denote the sum of the outcomes in each of the 36 rolls. Estimate the probability that $108 \leq Y \leq 144$.

40. A certain basketball player makes 80 percent of his free throws on average. What is the probability that in 100 attempts he will be successful more than 85 times? (Assume independence).

41. The number of students enrolled in a calculus class is a Poisson random variable with parameter 100. If the number enrolling is 120 or more, the class will be taught in two sections, otherwise it will be taught in one section. Find the probability that the class has to be taught in two sections.

42. Let X be a normal random variable with mean 0 and variance $a > 0$. What is the value of $P(X^2 < a)$?

43. An airline sells 200 tickets for a certain flight on an airplane that has 198 seats because, on average, 1% of purchasers of airline tickets do not appear for departure of their flight. Estimate the probability that everyone who appears for the departure of this flight will have a seat. (**Hint:** Use the law of rare events).