Part 2

Probability models

2.1 Exercises

1. A factory has a sensor measuring the level of noise pollution at 50 different random times daily. The noise level in each measurement can be modeled as a random variable whose density function is

$$f(x) = \begin{cases} ke^{-kx} & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

   (a) Determine the value of k, taking into account that the average noise pollution level is 26.3.

   (b) A noise pollution level above 60 is considered harmful. Determine the probability that on a given measurement the noise level goes above this threshold.

2. The daily proportion of successful requests to some SFTP is a random variable X with density function:

$$f(x) = \begin{cases} cx + d & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

   (a) Determine the values c and d for which the expectation of X is equal to 2/3.

   (b) What is the probability of having more failed than successful requests in a day?

3. A factory produces power plug-ins for oil refinement. The hexagonal interior of the plug-ins is made by means of electrolysis, and the diameter of this hexagon is of the utmost importance. We assume that this diameter follows a normal distribution with mean 20mm and standard deviation 0.01mm. On the other hand, the specification limits (the values the diameter should be between to be considered in order) for this diameter are 20.015 ± 0.025mm.

   (a) Determine the percentage of plug-ins that satisfy these specifications.

(b) After some adjustments, the variance of the diameter has remained the same, but the mean has increased to 20.015mm. What is now the percentage of plug-ins that satisfy the specifications?

(c) Taking these results into account, have the adjustments been effective?

4. A company that produces and bottles apple juice has a machine to fill up the half-liter bottles. However, there is some variation in the amount of liquid getting to each bottle. Previous studies have shown that the average amount of liquid is half a liter, with a standard deviation of 5 centiliters. Assuming that the amount of liquid getting to the bottle follows a normal distribution,

   (a) What is the probability of the machine pouring more than 550 ml in a bottle?

   (b) What is the amount of liquid for which only 5% of the bottles get more than this amount?

   (c) If the amount of liquid poured by another machine also follows a normal distribution with the same standard deviation as the former one and with unknown mean, what is the probability that the amount of liquid poured by the machine exceeds in 7.5 cl the mean?,

5. A football stadium has capacity for 29000 people. The number of people coming each weekend follows a distribution with mean 20000 and standard deviation 3000.

   (a) What is the probability that the average number of spectators, after 38 weekends, is greater than 21000?

   (b) If the number of spectators follows a normal distribution, what is the probability that on a given weekend the stadium is full?

6. The number of calls to the emergency number per day in a city is a random variable with mean 12 and standard deviation 6. Assuming that the distribution on different days are independent and identically distributed, what is the probability that in a month of 30 days there are more than 300 calls?

7. The lifetime of an electric component follows an exponential distribution with mean equal to 7 years.

   (a) What is the percentage of components that last longer than 14 years?

   (b) The manufacturing company specifies in its contract that it shall give the price of the component back if it lasts less than the guarantee issued. How long should this guarantee last if they want to return the money in at most 30% of the cases?

8. The lifetime, in years, of a component is a random variable X following an exponential distribution with parameter 0.5.

   (a) Calculate the average lifetime of these components.

(b) Determine the probability of a component lasting longer than this mean.

(c) The hazard rate of a component at time x ($h_X(x)$) is given by:

$$h_X(x) = \frac{f_X(x)}{1 - F_X(x)},$$

where f_X and F_X represent the density and distribution functions of X, respectively. Determine the hazard rate at the third year.

(d) A component is replaced immediately in case of failure, or after 4 years of use. What is the percentage of components that are replaced before they fail?

9. The lifetime (in days) X of a certain component follows a Weibull distribution with parameters $k = 0.5$ and $\lambda = 0.01$.

   (a) Determine the hazard function of X, $h_X(t) = \frac{f_X(t)}{1-F_X(t)}, t \in (0, \infty)$.

   (b) Taking this into account, may X model the lifetime of a component that deteriorates with time?

10. The lifetime in years of a mainframe follows Weibull distribution with parameters $k = 2$ and $\lambda = 3$.

    (a) Compute the probability of a mainframe lasting less than six years.

    (b) Estimate the reliability of the mainframe after three years, if the reliability at a given time is the probability of the mainframe lasting longer than this time.

    (c) Determine the quantile 0.75.

11. An urn has two white balls and three black ones. The draw of a white ball is rewarded with 10 euros and the black ball with 5 euros. We draw three balls without replacement from the urn.

    (a) Determine the expected reward.

    (b) If the draw of a black ball had a negative reward, how much should it be for the game to be fair?

12. We want to bet at the roulette at a casino, where the roulette has the following distribution of numbers:

Red/odd/1-12	Red/even/1-12	Red/odd/13-24	Red/even/13-24	Red/odd/25-36	Red/even/25-36
5	1	3	3	2	4

We consider a triple bet, to the color, dozen, and number (odd/even), and bet on the combination red/odd/1-12. The gains per euro are:

Combination	Triple	Double	Single	Nothing
Gain	4	1	0.5	0

(a) Determine the expected profit if we bet 10 euros.

(b) Determine the probability of winning something if we play only once.

(c) We bet 10 euros twice. Determine the probability of winning something exactly once.

13. A discrete random variable takes the values 0,1 and 3, and has the following distribution function:

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 0.4 & \text{if } 0 \leq x < 1 \\ 0.75 & \text{if } 1 \leq x < 3 \\ 1 & \text{if } x \geq 3. \end{cases}$$

Determine the probability mass function, the expectation and the variance.

14. Alice wants to go this weekend to watch the movie 'Fifty shades (really really) darker', while her boyfriend Bob prefers to watch 'Cannibal holocaust: the director's cut'. They decide to leave the choice to the highest card, so that each of them draws (with replacement) three cards from a poker pack. If Alice's highest card is a Queen, what is the probability that Bob wins the game? How many cards should Bob draw in order to have probability greater than 0.5 of winning the game? Assume that in case of draw Alice wins, and that the ace is considered greater than the King.

15. In the end Alice and Bob went to watch 'Fifty shades (really really) darker'. The time in minutes between the opening credits and Bob falling asleep follows an exponential distribution with expectation 45 minutes. If the movie is 90 minutes long, what is the probability of Bob not falling asleep? And if he is awake after the first hour?

16. The viscosity of some fillers for planes is an important feature for measuring its quality. We may assume that the viscosity of a batch follows a normal distribution with mean 33 and variance 0.16. The manufacturer makes quality controls where a batch is picked up at random and its viscosity is measured. If it is outside the interval 33 ± 0.8 there is a warning, which indicates that some readjustments should be made in the production. If it is outside the interval 33 ± 1.2 an alert is issued, which means that production is stalled.

(a) What is the expected percentage of batches that generate a warning?

(b) What is the percentage of batches that generate a warning, but not an alert?

(c) If 1000 batches are revised every trimester, how many times do we expect the production to be stalled?

17. A sugar refinery has three plants, all of which receive black sugar in bulk. The amount of sugar that each of these plants can process follows an exponential distribution with mean 4 tm.

(a) What is the probability that a plant processes more than 4 tm on a given day?

(b) If the three plants work independently, what is the probability that exactly two of them process more than 4tm on a given day?

(c) For a given plant, how much sugar should it store so that its probability of running out of sugar to process is of 0.05?

18. A company makes a quality control test over the components it receives from a supplier. For every batch of 1000 components, the company takes a sample of 30. If there are at most two defective ones, they are replaced by good ones and the batch is accepted. Otherwise, the whole batch is examined and all the defective components are replaced. If any component has probability 0.05 of being defective, compute:

    (a) the probability of not having to inspect the whole batch;

    (b) the probability distribution of the random variable "number of components examined";

    (c) the expected number of components examined per batch.

19. A company has ordered a supplier a large number of electric components. To decide upon its acceptance, they have devised a quality control system. The components come in batches of size 40000, and the company wants to accept only those batches with less than 1% of defective components. Therefore, they select a sample of 2.5% of the components of the batch and they accept it when this sample has at most 2% of faulty components (they reject it with 21 or more and they accept it with 20 or less).

    (a) We call business risk the probability of rejecting a batch which has a percentage of defective components no greater than 1%. Calculate the business risk with the quality control system above.

    (b) If the components came in batches of 4000 and the quality control system consisted in selecting 100 components and to accept the batch if there are at most 2 defective ones, what would be now the business risk?

    (c) Compare the answer to both questions.

20. The odds for Andy Merry in a tennis match against Rafa Nadull are 1:3, meaning that Andy is supposed to have three times as much probability of losing than of winning.

    (a) We are offered a bet where we get 10 euros if Andy wins, and we must pay 3 euros if he loses. What is the expected gain?

    (b) If Andy and Rafa play 50 times, what is the probability that Andy wins at least 10 times?

21. The number of breakdowns in an information system is a random variable X following a Poisson distribution with parameter 2.

    (a) What is the expected number of breakdowns?

(b) What is the probability that the number of breakdowns is between 1.5 and 3.7?

(c) What is the probability that the number of breakdowns lies within the interval $\mu \pm 3\sigma$, where μ and σ represent the mean and the standard deviation of X?

22. The number of incoming messages in a channel during an interval of t seconds follows a Poisson distribution with parameter $0.3\,t$. Determine the probability of the following events:

    (a) There are exactly two incoming messages in a 10 second period.

    (b) The number of incoming messages in a 5-seconds interval is between 2 and 4 (both included).

    (c) The time between two consecutive messages is longer than 4 seconds.

23. The number of arrivals per hour to a service follows a Poisson distribution $\mathcal{P}(10)$.

    (a) What is the probability that the average number of arrivals in the next 100 hours is between 9 and 11 (both included)?

    (b) What is the probability that more than 10 minutes pass between two consecutive arrivals?

24. The time it takes to an inhabitant of Gossipland to tell his closest neighbor a secret follows a exponential distribution with mean 0.8 days.

    (a) If Gossipland has 40 houses arranged in line and we have just told the inhabitant of the first house a secret, what is the probability that the inhabitant of the last house knows the secret before 30 days?

    (b) What is the probability that in a day two more people become aware of the secret?

2.2 Test

1. In a continuous probability distribution, what is the probability $P(X = \mu)$ of the random variable taking the value of its mean?

    (a) 0

    (b) 0.5

    (c) μ

    (d) 1

2. Which of the following conditions is NOT necessary for applying the Central Limit Theorem?

    (a) That the variables are independent.

(b) That they are continuous.

(c) That they are identically distributed.

(d) To have at least 30 variables.

3. Which of the following distributions can model the lifetime of a component that deteriorates in time?

(a) Weibull(1,2)

(b) Weibull(2,2)

(c) Weibull(0.5,2)

(d) exp(2)

4. Assume X follows a normal distribution with mean $\mu = 56$, and that 97.72% of its values are above 48. Then the standard deviation of X is:

(a) 8

(b) 1

(c) 4

(d) 50

5. The lifetime of a component follows an exponential distribution. The probability of a component lasting less than the average lifetime is:

(a) 1

(b) $1 - e^{-1}$

(c) e^{λ}/λ

(d) 97/3

6. Which of the following is NOT a feature of the binomial distribution?

(a) Each trial has a finite number of possible outcomes, and not necessarily 2.

(b) There is a fixed number of trials.

(c) The probability of success is the same for each of the trials.

(d) The trials are independent.

7. Consider eight blood donors selected at random. The probability of a blood donor being of the blood type A is 0.40. Which of the following is correct?

(a) The probability that at most 1 of the 8 donors is of type A is approximately 0.11.

(b) The probability of 7 or more donors NOT being of type A is approximately 0.0086.

(c) The probability of exactly 5 donors being of type A is approximately 0.28.

(d) The probability of exactly 5 donors being of type A is approximately 0.12.

8. Assume 10% of the components produced in a week where the process was maladjusted have some kind of defect. Consider the random variable $Y =$ "number of valid components in a sample of size 5 taken during that week". The probability distribution of Y is:

   (a) The same of the random variable $X =$ "number of defective components in a sample of size 5".

   (b) Binomial $B(5, 0.9)$.

   (c) Binomial $B(5, 0.1)$.

   (d) Poisson $\mathcal{P}(0.5)$.

9. The number of particles emitted by a radioactive source during a specific period is a random variable that follows a Poisson distribution. If the probability of having no emissions during the period is 1/3, then:

   (a) The expected number of particles emitted is $\ln(3)$.

   (b) The probability of having exactly one emission is $\ln(3)/3$.

   (c) The variance of the number of emissions during the period is $\ln(3)$.

   (d) The variance of the number of emissions during the period is $\ln(3)/3$.

10. If the number of cars arriving to a garage per day follows a Poisson distribution with variance equal to 2, then the probability of no cars arriving for a day is :

    (a) 0.

    (b) e^{-2}.

    (c) $e^{-\sqrt{2}}$.

    (d) $1 - e^{-\sqrt{2}}$.

11. Which of these properties MUST be satisfied by a density function?

    (a) It takes values in $[0, 1]$.

    (b) It is an increasing function.

    (c) The integral between $-\infty$ and $+\infty$ is 1.

    (d) Its limit as we go to $+\infty$ is 1.

12. We roll a fair dice, so that each of the results 1,2,3,4,5,6 has probability 1/6. What is the expectation of the outcome?

    (a) 3.5.

    (b) 7.

(c) 2.12.

(d) 3.

13. Assume $X \sim N(3,2)$. How much is $P(X > 4)$?

(a) 0.3085.

(b) 0.6915.

(c) 0.

(d) 0.2514.

14. To which distribution can we approximate a binomial $B(50, 0.3)$?

(a) Poisson $\mathcal{P}(15)$.

(b) Normal $N(15, 3.24)$.

(c) Normal $N(0.3, 0.0648)$.

(d) Normal $N(15, 10.5)$.

15. If the number of arrivals to a shop per hour follows a Poisson distribution $\mathcal{P}(10)$, what is the distribution of the time in minutes between consecutive arrivals?

(a) Poisson $\mathcal{P}(1/6)$.

(b) exponential $exp(10)$.

(c) exponential $exp(600)$.

(d) exponential $exp(1/6)$.

2.3 Solutions

1. (a) f is the density function of a exponential distribution with parameter $k > 0$. The expectation of an exponential distribution is the inverse of this parameter, so in this case $\int_{-\infty}^{\infty} x\, f(x)\, dx = 1/k$. Hence, the condition is satisfied if k is $k = 1/26.3 \approx 0.038$.
   Alternatively, from the equation $26.3 = E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^{+\infty} kxe^{-kx} dx$ we deduce, applying integration by parts, that it must be $k = \frac{1}{26.3}$.

   (b) If X denotes the random variable "noise pollution level", we must compute $P(X > 60) = 1 - P(X \leq 60) = 1 - F(60) = 1 - (1 - e^{-0.038 \cdot 60}) \approx 0.1$.

2. (a) In order to have $1 = \int_{-\infty}^{+\infty} f(x)dx$ and $E(X) = \int_{-\infty}^{+\infty} xf(x)dx = 2/3$, it must be $c = 2$ and $d = 0$.

   (b) $P(X < 1 - X) = P(X < 0.5) = 1/4$

3. (a) $P(-1 \leq N(0,1) \leq 4) = 0.8413\%$.

(b) $P(-2.5 \leq N(0,1) \leq 2.5) = 0.9876\%$.

(c) Yes, because from a percentage of non-conforming components of almost 16% we have come to less than 2%, which means an increase in the quality of the production.

4. Let X denote the random variable that represents the amount of juice (in cl.) poured in a bottle selected at random. This variable follows a distribution N(50, 5).

   (a) The probability of pouring more than 55 cl is $1 - F_X(55) = 1 - 0.8413 = 0.1587$.

   (b) We seek the value c such that $P(X > c) = 0.05$, or, equivalently, $P(X \leq c) = 0.95$. Hence, $c = 58.224$ cl.

   (c) Let Y denote the random variable representing the amount of juice (in cl.) poured in a bottle selected at random with the new machine. It follows a distribution $N(\mu, 5)$, where μ is unknown.

   We must compute $P(X > \mu + 7.5)$, which is equal to:

$$P(X > \mu + 7.5) = 1 - P(X \leq \mu + 7.5) = 1 - F(\mu + 7.5) =$$
$$1 - \phi\left(\frac{(\mu+7.5)-\mu}{5}\right) = 1 - \phi(1.5) = 1 - 0.9332 = 0.0668.$$

   Hence, we expect a 6.68% of pours that exceed the mean in 7.5 cl.

5. Let X_i be the random variable "number of spectators on the i-th game", $E(X_i) = 20000, SD(X_i) = 3000$.

   (a) The mean number of spectators is $X \sim \frac{X_1+\cdots+X_{38}}{38}$. Using the Central Limit Theorem, we can approximate $X \sim N(20000, \frac{3000}{\sqrt{38}}) = N(20000, 486.6)$.

$$P(X > 21000) = P(N(0,1) > 2.05) = 0.021.$$

   (b) $P(N(20000, 3000) > 29000) = P(N(0,1) > 3) = 0.0014$.

6. Let X_i denote the random variable 'number of calls on the i-th day'. It follows that $E(X_i) = 12, SD(X_i) = 6$. Since the variables are assumed to be independent and identically distributed, we can apply the Central Limit Theorem and approximate

$$Z = X_1 + \cdots + X_{30} \sim N(30 \cdot \mu, \sqrt{30}\sigma) = N(360, 32.86).$$

   Thus, $P(Z > 300) = P(N(360, 32.86) > 300) = P(N(0,1) > -1.83) = P(N(0,1) < 1.83) = 0.9664$.

7. (a) Since $X \equiv exp(1/7)$, we have that

$$P(X > 14) = 1 - P(X \leq 14) = 1 - F(14) = 1 - (1 - e^{-2}) = 0.1353.$$

   Hence, the expected percentage of components lasting more than 14 years 13.53%.

(b) We are seeking for the time t such that $P(X \leq t) = 0.3$. Hence,

$$0.3 = P(X \leq t) = F(t) = 1 - e^{-t/7}$$

This equality is equivalent to $e^{-t/7} = 0.7 \Leftrightarrow t = -7ln(0.7) = 2.5$, so the guarantee should be of two and a half years maximum.

8. (a) 2 years.
   (b) $P(X > 2) = 0.368$.
   (c) Since $h_X(x) = \frac{f_X(x)}{1-F_X(x)} = \frac{0.5e^{-0.5x}}{e^{-0.5x}} = 0.5$, the hazard rate is constant and equal to 0.5.
   (d) Since $P(X < 4) = 0.8647$, we expect that 13.53% are replaced without having failed.

9. The density function of X is $f_X(x) = 5x^{-0.5}e^{-10x^{0.5}}$, when $x > 0$, $f(x) = 0$, if $x \leq 0$. Its associated distribution function is $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt,\ \forall\, x$, so $F_X(x) = 1 - e^{-10\,x^{0.5}}$, if $x > 0$ and $F_X(x) = 0$, if $x \leq 0$.

   (a) $h(t) = \frac{f(t)}{1-F(t)} = \frac{5\,t^{-0.5}\,e^{-10\,t^{0.5}}}{e^{-10\,t^{0.5}}} = 5\,t^{-0.5},\ \forall t > 0$
   (b) The function h is decreasing, so X cannot represent the lifetime of a component that deteriorates with time.

10. If $X \equiv W(2,3)$, then

$$P(X \leq x) = F(x) = 1 - e^{-(x/3)^2},\ \text{si } x \geq 0.$$

   (a) The probability of failure before 6 years is 98.17%, because

$$P(X < 6) = P(X \leq 6) = F(6) = 1 - e^{-(6/3)^2} = 1 - e^{-4} \simeq 0.9817.$$

   (b) The reliability after 3 years is 36.79%, because

$$P(X > 3) = 1 - P(X \leq 3) = 1 - F(3) = 1 - (1 - e^{-(3/3)^2}) = e^{-1} \simeq 0.3679.$$

   (c) The quantile 0.75 of this distribution would be the value k such that $P(W(2,3) \leq k) = 0.75$. Using the formula of the distribution function of the Weibull, we obtain that

$$0.75 = P(W(2,3) \leq k) = 1 - e^{-(k/3)^2} \Rightarrow e^{-(\frac{k}{3})^2} = 0.25 \Rightarrow \left(\frac{k}{3}\right)^2 = 1.38 \Rightarrow k = 3.53.$$

11. (a) Let us consider the random variables X="number of white balls drawn" and Y="earnings". They have the following probability distribution:

Y	X	Prob.
15	0	0.1
20	1	0.6
25	2	0.3

To obtain these probabilities, use for instance that $P(X=0) = P(BBB) = \frac{3}{5}\cdot\frac{2}{4}\cdot\frac{1}{3} = 0.1$, and similarly for the other combinations.

As a consequence, $E(Y) = 21$.

(b) If the draw of a black ball is penalized with a loss of x euros, the probability distribution of Y becomes

Y	Prob.
-3x	0.1
10-2x	0.6
20-x	0.3

For the game to be fair, we should have $E(Y) = 12 - 1.8x = 0$, so $x = 6.6$

12. The probabilities of the different gains are:

Combination	Gain	Probability
Triple	30	$\frac{5}{37}$
Color and dozen	0	$\frac{1}{37}$
Color and number	0	$\frac{5}{37}$
Dozen and number	0	$\frac{1}{37}$
Color	-5	$\frac{7}{37}$
Dozen	-5	$\frac{5}{37}$
Number	-5	$\frac{7}{37}$
Nothing	-10	$\frac{6}{37}$

where we are subtracting the 10 euros we bet from each of the rewards. Remember also that we must take the number 0 into account.

(a) The expected profit is

$$E(X) = 30\cdot\frac{5}{37} - 5\cdot\frac{19}{37} - 10\cdot\frac{6}{37} = -\frac{5}{37}.$$

(b) If we play once, the probability of winning some money is $\frac{5}{37}$.

(c) If we play twice, the probability of winning money exactly once is $P(B(2,\frac{5}{37}) = 1) = 0.233$.

13. From the values of the distribution function, we deduce that $P(X=0) = F(0) - F(0^-) = 0.4, P(X=1) = F(1) - F(1^-) = 0.35$ and $P(X=3) = F(3) - F(3^-) = 0.25$. From this it follows that the expectation is

$$E(X) = 0\cdot 0.4 + 1\cdot 0.35 + 3\cdot 0.25 = 1.1,$$

and that the variance is

$$Var(X) = E((X-E(X))^2) = (0-1.1)^2\cdot 0.4 + (1-1.1)^2\cdot 0.35 + (3-1.1)^2\cdot 0.25 = 1.39.$$

14. In every draw, Bob has probability 0.2 of drawing a King or an ace. His probability of winning the bet is thus

$$P(B(3, 0.2) \neq 0) = 1 - 0.8^3 = 0.488.$$

If he drew n cards, his probability of winning would be

$$P(B(n, 0.2) \neq 0) = 1 - 0.8^n,$$

which is greater than 0.5 if and only if

$$0.8^n \leq 0.5 \Leftrightarrow n \geq \frac{log(0.5)}{log(0.8)} = 3.1,$$

that is, from the fourth draw.

15. Let X be the random variable that tells us the number of minutes before Bob falls asleep, $X \sim exp(\frac{1}{45})$. Then,

$$P(X > 90) = e^{-2} = 0.135,$$

so $P(X > 90|X > 60) = P(X > 30) = e^{-\frac{2}{3}} = 0.513$.

16. (a) We expect 4.55% of warnings, because $P(32.2 \leq X \leq 33.8) = P(X \leq 33.8) - P(X < 32.2) = 0.9773 - 0.0228 = 0.9545$.

    (b) Since the probability of not issuing an alert is $P(31.8 \leq X \leq 34.2) = P(X \leq 34.2) - P(X < 31.8) = 0.9987 - 0.0013 = 0.9974$ and every alert is in particular under the conditions of a warning, the expected percentage of batches generating a warning but not an alert is $4.55 - 0.26 = 4.29\%$.

    (c) Let Y denote the random variable "number of standstill produced by the inspection of 1000 batches". This variable follows a binomial distribution $B(1000, 0.0026)$. Since we must determine the mean of Y and for a binomial $B(n, p)$ this is equal to $n \cdot p$, we conclude that the expected number of standstill is $1000 \cdot 0.0026 = 2.6$.

17. (a) The probability that a plant processes more than 4tm in a day is equal to the probability of an exponential distribution exp(1/4) being greater than 4. This probability is equal to e^{-1}.

    (b) Let Y denote the random variable representing the number of plants that process more than 4 tons in a day. It follows a binomial distribution with parameters $n = 3$ and $p = e^{-1}$, so $P(Y = 2) = 3\,e^{-2}\,(1 - e^{-1}) = 0.255$.

    (c) We must store 11.982 tons, because $P(X > 11.982) = 0.05$, where X represents the amount of sugar processed by each plant in a day.

18. (a) If $X \equiv B(30, 0.05)$ then

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) =$$

$$\binom{30}{0}(0.05)^0(0.95)^{30} + \binom{30}{1}(0.05)^1(0.95)^{29} + \binom{30}{2}(0.05)^2(0.95)^{28} =$$

$$0.2146 + 0.3389 + 0.2586 = 0.8121.$$

(b) We check 30 components when the batch passes the test and 1000 when it does not, so the probability distribution of the random variable Y = "number of components checked" is

Y	30	1000
P_Y	0.8121	0.1879

(c) The expected number of checked components is 212.263.

19. (a) Let X denote the number of defective components in a sample of size 1000. This random variable follows a distribution $B(1000, p)$, where p is the probability that each component of the batch has of being defective. The business risk is the probability $P(X > 20)$ when $p = 0.01$, i.e., $P(B(1000, 0.01) > 20)$. Using the approximation by a normal distribution from the Central Limit Theorem, it follows that $B(1000, 0.1) \approx N(10, \sqrt{9.9})$, so $P(B(1000, 0.1) > 20) \approx P(N(10, \sqrt{9.9}) > 20) = 0.0008$.

(b) Now X follows a binomial distribution $B(100, 0.01)$, which we can approximate by a normal distribution $N(1, \sqrt{0.99})$ and we reject the batch when $X > 2$. The business risk now is

$$P(B(100, 0.01) > 2) \approx P(N(1, \sqrt{0.99}) > 2) = 0.157.$$

(c) In both cases we have selected a sample of 2.5% of the components of the batch (1000 out of 40000 and 100 out of 4000, respectively) and we accept the batch when at most 2% of the selected components are defective (20 out of 1000 and 2 out of 100, respectively). Hence, both systems seem to be equivalent. However, their business risks are different, which means that the quality control system should take into account not only the % of inspected batches, but also the sample size n.

20. According to these odds, if p is the probability that Andy has of winning the match, it holds that $1 - p = 3p$, so $p = 0.25$.

(a) The random variable X that provides the gain has the following probability distribution:

x	p(x)
10	0.25
-3	0.75

Thus, the expected gain is $E(X) = 10 \cdot 0.25 - 3 \cdot 0.75 = 0.25$ euros.

(b) Let Y be the random variable 'number of times Andy wins in 50 matches'. Then, Y follows a binomial distribution $B(50, 0.25)$ that we can approximate by a normal distribution $N(50 \cdot 0.25, \sqrt{50 \cdot 0.25 \cdot 0.75}) = N(12.5, 3.06)$ using the Central Limit Theorem. As a consequence,

$$\begin{aligned} P(Y \geq 10) \approx P(N(12.5, 3.06) \geq 10) &= P(N(0,1) \geq -0.81) \\ &= P(N(0,1) \leq 0.81) = 0.7910. \end{aligned}$$

21. (a) $E(X) = 2$.

(b) $P(1.5 < X < 3.7) = P(X = 2) + P(X = 3) = e^{-2}\frac{2^2}{2!} + e^{-2}\frac{2^3}{3!} = 0.2707 + 0.1804 = 0.4511$.

(c) Since $\mu - 3\sigma = 2 - 3\sqrt{2} = -2.24$ and $\mu + 3\sigma = 2 + 3\sqrt{2} = 6.24$, we must compute $P(-2.24 \le X \le 6.24)$, which is equal to

$$\begin{array}{l} P(X=0)+P(X=1)+P(X=2)+P(X=3)+ \\ P(X=4)+P(X=5)+P(X=6)=0.9955 \end{array}$$

22. (a) $e^{-3}\,\frac{3^2}{2!} = 4.5\,e^{-3} = 0.224$.

(b) $e^{-1.5}(\sum_{i=2}^{4} \frac{1.5^i}{i!}) = 0.424$.

(c) $e^{-0.3\,4} = 0.301$.

23. (a) Let X_i denote the number of arrivals in the i-th hour, for $i = 1, \ldots, 100$. Then given $Z = \frac{X_1+\cdots+X_{100}}{100}$, we must compute $P(9 \le Z \le 11)$.
Applying the Central Limit Theorem, we can approximate the distribution of Z by a normal $N(10, \sqrt{\frac{10}{100}})$=N(10,0.316). Hence,

$$P(9 \le Z \le 11) = P(-3.16 \le N(0,1) \le 3.16) = 0.9992 - 0.0008 = 0.9984.$$

(b) The time in hours between two consecutive arrivals is a random variable Y that follows an exponential distribution $exp(10)$. From this we deduce that

$$\begin{aligned} &P(\text{ more than 10 minutes between consecutive arrivals}) \\ &\qquad = P\left(Y > \frac{1}{6}\right) = e^{-10\cdot\frac{1}{6}} = 0.1887. \end{aligned}$$

24. (a) Let X_i="days it takes the i-th neighbor to tell the $i+1$-th the secret". Then $X_i \sim \exp(\frac{1}{0.8})$, and the total number of days is $X_1 + \cdots + X_{39} \sim N(31.2, 5)$, using the Central Limit Theorem. Hence,

$$P(N(31.2, 5) < 30) = 0.4052$$

(b) The number of people that become aware of the secret on a given day follows a Poisson distribution $\mathcal{P}(\frac{1}{0.8})$, so

$$P(\mathcal{P}(1.25) = 2) = e^{-1.25}\frac{1.25^2}{2} = 0.223$$

2.4 Test solutions

1. a)
2. b)
3. b)
4. c)
5. b)
6. a)
7. a) and d)
8. b)
9. a), b) and c)
10. b)
11. c)
12. a)
13. a)
14. b)
15. d)