University of Geneva   Bachelor in Economics and Management
Probability 1   Bachelor in International Relations
Professor Davide La Vecchia   Spring 2018
ASSIGNMENT 06   April 3rd

Exercise 1

We consider an urn with N marbles containing r red marbles and $N - r$ blue marbles. We draw n marbles randomly without replacement. Let X be the following discrete random variables:

$X = $ "the number of red marbles among the n drawn marbles".

1. Compute the probability law of X, that is $P(X = k)$ for $k = 0, 1, ..., min(n, r)$.

2. This law is well known. Its expectation is $E[X] = np$ and its variance $var(X) = np(1 - p)\frac{N-n}{N-1}$ with $p = \frac{r}{N}$. Compare intuitively this law with the binomial law.

Solutions:

1. X ={the number of red marbles among the n drawn marbles}
   S = {draw n marbles randomly without repetition from N marbles}

$$P(X = k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}, k = 0, 1, ..., \min(n, r).$$

2. Let $Y \sim Bin(n, p)$. We can see a binomial distribution as a drawing with replacement where $p = \frac{r}{N}$. We know that $E(X) = np$ and $V(Y) = np(1 - p)$. So:

$$\begin{aligned} E(X) &= E(Y) \\ V(X) &= \frac{N-n}{N-1}V(Y) \le V(Y). \end{aligned}$$

If r and N increase to infinity while the ratio $p = r/N$ is constant, they are equivalent.

Exercise 2

An editor is interested to model the number of mistakes in his books. We consider a book of $n = 200$ pages. The number of typos for each page follows a Poisson distribution with parameter $\lambda = 0.05$ and is independent of the number of typos on the other pages.

Let us do some preliminary considerations. Define

$$X_i = \{\text{the number of typos on the } i_{th} \text{ page}\} \quad i = 1, 2, ..., n$$

Each X_i is independent and follows a Poisson distribution with parameter $\lambda = 0.05$. On page i-th, the probability of no typos is

$$P(X_i = 0) = \frac{\lambda^0 e^{-\lambda}}{0!} = e^{-\lambda} = e^{-0.05}$$

1. What is the expected number of pages without typos?
   Hint: define a random variable Y_n which represents the number of pages without any typo and compute $E(Y_n)$.

   $Y_n = \{$the number of pages without any typos$\}$. Y_n follows a Binomial distribution with parameter $n = 200$ and $p = P(X_i = 0)$. The expected value of Y_n is $np = 200e^{-0.05}$.

2. What is the variance of the number of pages without any typo?

   The variance of Y_n is $np(1-p) = 200e^{-0.05}(1 - e^{-0.05})$.

3. What is, approximately, the probability that a book has at least 195 pages without typos?

   The probability that a book has at least 195 pages without typos is $P(Y_n \geq 195)$. we know p is large, so we need to transform Y_n to get a small p. Then we define Z_n: $\{$the number of pages with typos$\}$. $Z_n = 200 - Y_n$. Z_n also follows a binomial distribution with parameter n and $q = 1 - p \approx 0.05$. As n is relatively large and q is relatively small, we can use Poisson to approximate the required probability. So,

$$\lambda = nq = 200(1 - e^{-0.05})$$
$$P(Y_n \geq 195) = P(Z_n \leq 5) \approx \sum_{i=0}^{5} \lambda^{-i} \frac{e^{-\lambda}}{i!} \approx 0.077$$

Exercise 3

Charles-Basile passes a test of probabilities and statistics. The test is a multiple choice quiz with 6 questions. Each question has three possible answers, only one of which is correct. Charles-Basile passes the test if he answers correctly at least 4 questions.

1. If Charles-Basile answers at random, what is the expectation of the number of correct answers? What is its variance? What is the probability that Charles-Basile will pass the test?

   Let R be the random variable of the number of correct answers (number of successes). If Charles-Basile randomly answers a question, as there are three possibilities, he

has a probability of $\frac{1}{3}$ to choose the correct answer. In addition, there are 6 questions in all and they are independent. Thus, R is a binomial random variable with parameters $n = 6$ and $p = \frac{1}{3}$:

$$E(R) = np = 6 \cdot \frac{1}{3} = 2.$$

The variance of the number of correct answers is given by:

$$\text{var}(R) = np(1 - p) = 6 \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{4}{3}.$$

To pass the test, Charles-Basile must answer correctly at least 4 questions, *i.e.* $P(\text{pass the test}) = P(R \geq 4)$. We calculate the probability that Charles-Basile will pass the test using the probability function of R:

$$
\begin{aligned}
P(R \geq 4) &= P(R = 4) + P(R = 5) + P(R = 6) \\
&= \binom{n}{4}p^4(1-p)^{n-4} + \binom{n}{5}p^5(1-p)^{n-5} + \binom{n}{6}(1-p)^{n-6} \\
&= \frac{6!}{4!2!}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^2 + \frac{6!}{5!1!}\left(\frac{1}{3}\right)^5\left(\frac{2}{3}\right)^1 + \frac{6!}{6!0!}\left(\frac{1}{3}\right)^6\left(\frac{2}{3}\right)^0 \\
&= \frac{60}{729} + \frac{12}{729} + \frac{1}{729} \\
&= \frac{73}{729} \approx 0.10014.
\end{aligned}
$$

2. Being a little better prepared, Charles-Basile is able to eliminate an incorrect answer for each question. Then he chooses his answer randomly from the two remaining possibilities. Find the expectation and variance of the number of correct answers in this case. What is the probability that Charles-Basile will pass the exam?

The only difference from the first case is that now, Charles-Basile is able to eliminate a false answer. As a result, the probability of success for a given question becomes $\frac{1}{2}$. Let $\tilde{R}$ be the random variable of the number of correct answers (number of successes) when Charles-Basile is able to eliminate a false answer. There are $n = 6$ independent questions, but now $\tilde{p} = \frac{1}{2}$. We then have:

$$\tilde{R} \sim B\left(6, \frac{1}{2}\right).$$

The expectation of the number of correct answers is given by the expectation of $\tilde{R}$, which is calculated as before:

$$E(\tilde{R}) = n\tilde{p} = 6 \cdot \frac{1}{2} = 3.$$

3

The variance of the number of correct answers is given by:

$$\text{var}(\tilde{R}) = n\tilde{p}(1 - \tilde{p}) = 6 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{2}.$$

As in the previous part, we must calculate $P(\tilde{R} \geq 4)$, which corresponds to the probability that Charles-Basile will pass the test:

$$
\begin{aligned}
P(\tilde{R} \geq 4) &= P(\tilde{R} = 4) + P(\tilde{R} = 5) + P(\tilde{R} = 6) \\
&= \binom{n}{4}\tilde{p}^4(1 - \tilde{p})^{n-4} + \binom{n}{5}\tilde{p}^5(1 - \tilde{p})^{n-5} + \binom{n}{6}\tilde{p}^6(1 - \tilde{p})^{n-6} \\
&= \frac{6!}{4!2!}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^2 + \frac{6!}{5!1!}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^1 + \frac{6!}{6!0!}\left(\frac{1}{2}\right)^6\left(\frac{1}{2}\right)^0 \\
&= \frac{15 + 6 + 1}{64} \\
&= \frac{22}{64} = 0.34375.
\end{aligned}
$$

Exercise 4

The phone calls to 'Residence Soleil' follow the Poisson distribution with $\lambda = 2$ calls per hour.

Let R be the random variable representing the number of phone calls per hour. We hypothesize: $R \sim \text{Poisson}(\lambda)$, with $\lambda = 2$.

1. What is the average number of calls in this residence? What is the variance of the number of calls?

   The average number of calls per hour is given by the expectation of R. Using the properties of Poisson's law, we have:

   $$\text{E}(R) = \lambda = 2.$$

   We can also find the variance by the properties of the Poisson's law:

   $$\text{var}(R) = \lambda = 2.$$

2. Calculate the probability of receiving more than two calls in one hour.

The probability of receiving more than two calls in one hour is calculated as follows:

$$
\begin{aligned}
P(R > 2) &= 1 - P(R \leq 2) \\
&= 1 - [P(R=0) + P(R=1) + P(R=2)] \\
&= 1 - P(R=0) - P(R=1) - P(R=2) \\
&= 1 - \frac{\lambda^0}{0!}\exp(-\lambda) - \frac{\lambda^1}{1!}\exp(-\lambda) - \frac{\lambda^2}{2!}\exp(-\lambda) \\
&= 1 - \exp(-2)(1 + 2 + 2) \\
&= 1 - \frac{5}{\exp(2)} \approx 0.3233.
\end{aligned}
$$

3. Charles-Basile, living in 'Residence Soleil', would like to take a shower for 10 minutes. What is the probability that the phone will ring while Charles-Basile is in the shower?

Let $\tilde{R}$ be the random variable representing the number of calls received for 10 minutes. As 10 minutes$= \frac{1}{6}$ hour, we have that $\tilde{R}$ follows a Poisson's distribution with a parameter $\tilde{\lambda} = \frac{1}{6}\lambda = \frac{1}{3}$. The probability of the phone ringing in this 10 minutes is given by:

$$
\begin{aligned}
P(\tilde{R} > 0) &= 1 - P(\tilde{R} \leq 0) = 1 - P(\tilde{R} = 0) \\
&= 1 - \frac{\tilde{\lambda}^0}{0!}\exp(-\tilde{\lambda}) = 1 - \exp\left(-\frac{1}{3}\right) \\
&\approx 1 - 0.7165 = 0.2835.
\end{aligned}
$$

Exercise 5

1. In a large sports store, it is estimated that the probability that a randomly selected customer is a thief is 0.005. The store welcomes 400 customers daily and it is assumed that an individual does not engage in more than one theft a day.

Let R be the random variable of the number of thefts per day in a department store. $n = 400$ is the number of customers per day and $p = 0.005$ is the probability of 'success'. R follows a binomial distribution, $R \sim B(n,p)$. So, for $k \in \{0, ..., 400\}$:

$$
P(R = k) = \binom{n}{k}p^k(1-p)^{n-k}.
$$

Since n is large and p is small, we can use Poisson's Law as the approximation of the binomial distribution of R with $\lambda = np = 2$. In other words, $R \sim \text{Poisson}(\lambda)$ approximately. So, for $k \in \{0, ..., 400\}$:

$$
P(R = k) \approx \frac{\lambda^k}{k!}\exp(-\lambda).
$$

Calculate the probability that there are more than 2 thefts per day using

(a) Poisson distribution;

$$
\begin{aligned}
P(R > 2) &= 1 - P(R \le 2) = 1 - P(R = 0) - P(R = 1) - P(R = 2) \\
&\approx 1 - \frac{\lambda^0}{0!}\exp(-\lambda) - \frac{\lambda^1}{1!}\exp(-\lambda) - \frac{\lambda^2}{2!}\exp(-\lambda) \\
&= 1 - (1 + 2 + 2)\exp(-2) \approx 0.3233.
\end{aligned}
$$

(b) Binomial distribution.

$$
\begin{aligned}
P(R > 2) &= 1 - P(R \le 2) = 1 - P(R = 0) - P(R = 1) - P(R = 2) \\
&= 1 - \binom{n}{0}p^0(1 - p)^{n-0} - \binom{n}{1}p^1(1 - p)^{n-1} - \binom{n}{2}p^2(1 - p)^{n-2} \\
&= 1 - \frac{400!}{0!400!}(0.005)^0(0.995)^{400} - \frac{400!}{1!399!}(0.005)^1(0.995)^{399} \\
&\quad - \frac{400!}{2!398!}(0.005)^2(0.995)^{398} \approx 1 - 0.67668 = 0.32332.
\end{aligned}
$$

Compare the expected number of thefts per day in both cases. Comment on the results obtained.

The approximation of the binomial distribution with Poisson's law is good since n is large enough and p is small enough. The result is almost identical here.

In all cases, the expectation is $\lambda = np = 2$ (by construction).

2. It can be seen that 50% of thefts are for items at 50 CHF, 20% for items at 100 CHF, and 30% for items at 200 CHF. Hiring a supervisor would make thefts decrease by half and cost the store management 130 CHF per day.

Does the management have an interest in hiring a supervisor? To justify.

We put $R = R_1 + R_2 + R_3$, with:

$$
\begin{aligned}
R_1 &= \text{number of thefts on items at 50 CHF;} \\
R_2 &= \text{number of thefts on items at 100 CHF;} \\
R_3 &= \text{number of thefts on items at 200 CHF.}
\end{aligned}
$$

Let X be the random variable that corresponds to the loss caused by thefts. So:

$$
X = 50R_1 + 100R_2 + 200R_3.
$$

So the expectation of X is given by:

$$
\mathrm{E}(X) = 50\mathrm{E}(R_1) + 100\mathrm{E}(R_2) + 200\mathrm{E}(R_3).
$$

We have $R_i \sim B(n, p_i)$, with p_i the probability of having a theft of a value (in CHF) of v_i, $i = 1, 2, 3$ ($v_1 = 50, v_2 = 100, v_3 = 200$). We have:

$$
p_i = P(\text{value} = v_i|\text{thefts})P(\text{thefts}) = P(\text{value} = v_i|\text{thefts}) \cdot p, \quad i = 1, 2, 3.
$$

Therefore, $p_1 = 0.5p$, $p_2 = 0.2p$, $p_3 = 0.3p$, and:

$$\mathrm{E}(R_1) = 0.5np, \mathrm{E}(R_2) = 0.2np, \mathrm{E}(R_3) = 0.3np.$$

So

$$\mathrm{E}(X) = (50 \cdot 0.5 + 100 \cdot 0.2 + 200 \cdot 0.3)np = 105np.$$

We have $n = 400$.

If no supervisor is hired, we have $p = 0.005$, then $\mathrm{E}(X) = 105 \cdot 2 = 210$ CHF.

If a supervisor is hired, the probability of theft is reduced by half and thus goes to $\frac{0.005}{2} = 0.0025$. In that case, $\mathrm{E}(X) = 105 \cdot 1 = 105$ CHF.

The commitment of a supervisor reduces the expectation of daily loss of 105 CHF but costs 130 CHF/day (certain cost). If management relies on the expected loss to make the decision, they should not hire a supervisor since the expected decrease in the loss is less than his salary.

Exercise 6 (Optional)

Pamela has decided to do a trip of 10 000km this summer in a Fiat Ritmo. But the probability of Pamela having an accident during a 1km distance is 1/10 000. Knowing this probability, she decides to cancel her long trip arguing that she is sure to have an accident. Do you agree with Pamela? If it is not the case, what is her mistake? Compute approximately the probability of Pamela having an accident.

Solutions:

Pamela thinks: "I have a probability of $10000 \cdot \frac{1}{10000} = 1$ of having a accident". In fact, let's call X the random variable counting the number of accidents during the trip. It is a binomial with parameters $(10000, 0.0001)$ and we can approximate it by a poisson with parameter 1. $P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-1} \simeq 0.63$. She has "only" 63% chance of having an accident during the 10 000km trip.

Exercise 7 (Optional)

We consider a series of independent trials. At each trial, we observe a success with probability p and a failure with probability $1 - p$. Let X be the random variable counting the number of trials necessary to obtain the first success.

1. Compute the probability law of X, that is $P(X = k)$, for $k = 1, 2, ...$

2. Check that $E(X) = 1/p$.
   Hint: Use the formula of the geometric difference series: $\sum_{k=0}^{\infty} kq^{k-1} = \frac{1}{(1-q)^2}$.

3. Check that the variable is without memory, which means:

$$P(X > k | X > j) = P(X > k - j) \quad \text{for} \quad k > j.$$

4. I decided to sell my house and to accept the first purchase offer bigger than K CHF. We assume that the purchase offers are independently distributed random variables with known cumulative distribution F. Let N be number of purchase offers received before selling the house. Compute the probability law of N, that is $P(N = n)$, for $n = 1, 2, ...$

Solutions:

1. $P(X = k) = (1 - p)^{k-1}p$ for $k = 1, 2, \cdots$

2. Let $q = 1 - p$:

$$
\begin{aligned}
E(X) &= p \sum_{k=0}^{\infty} kq^{k-1} = p \sum_{k=0}^{\infty} \frac{d}{dq} q^k \\
&= p \frac{d}{dq} \sum_{k=0}^{\infty} q^k = p \frac{d}{dq} \frac{1}{1-q} \\
&= p \frac{1}{(1-q)^2} = \frac{1}{p}
\end{aligned}
$$

3. $P(X > k) = \sum_{i=k+1}^{\infty} q^i p = p \frac{q^{k+1}}{1-q} = (1-p)^{k+1}$
   Then: $P(X > k | X > j) = \frac{P(X>k))}{P(X>j)} = \frac{(1-p)^k}{(1-p)^j} = q^{k-j} = P(X > k - j)$

4. N follows an geometric law with parameter $P(X > K) = 1 - F(K)$. So: $P(N = n) = F(K)^{n-1}(1 - F(K))$.