Lectures 9–10. Discrete Probability Distributions.

YULIA R. GEL

CS/SE/STAT 3341 Probability and Statistics in Computer Science and Software Engineering

February 14 and 16, 2017

1. Bernoulli distribution
2. Binomial Distribution
3. Geometric Distribution
4. Poisson Distribution

Bernoulli trials

If an experiment involves a sequence of independent but identical stages, we say that we have a sequence of **independent trials**.

In the special case where there are only two possible results at each stage, i.e. a binary outcome, we say that we have a sequence of independent **Bernoulli trials**.

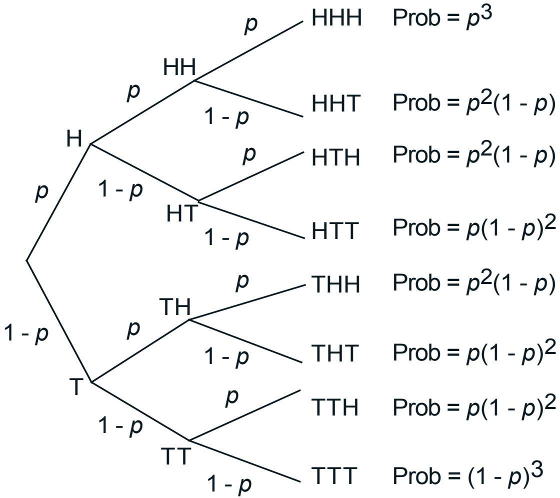
Examples. The two possible results can be anything, e.g., "pass" or fail, "head" or tail, optimisation routine has converged with bugs (0) vs. has not converged due to bugs (1).

Example: Coin

Let us have 3 Bernoulli trials of tossing a (not necessarily fair!) coin such that probability of getting a head is p.

Then, we can visualize it as follows (Bertsekas & Tsitsiklis, 2000):

Example: Coin (Bertsekas & Tsitsiklis, 2000)



Example: Coin

The conditional probability of any toss being H, conditioned on the results of any preceding tosses is p, because of independence.

Hence, by multiplying the conditional probabilities along the corresponding path of the tree, we see that any particular outcome (3-long sequence of H and T) that involves k heads and $3-k$ tails has probability $p^k(1-p)^{3-k}$.

Example: Coin

This formula extends to the case of a general number n of tosses. We obtain that the probability of any particular n-long sequence that contains k heads and $n-k$ tails is

$$p^k(1-p)^{n-k},$$

for all $0 \leq k \leq n$.

Bernoulli distribution

Definition. Formally, we say that a r.v. X is a **Bernoulli variable**, or follows **Bernoulli distribution**, if it has binary outcomes 0 or 1.

Properties. Like in the previous coin example, let us assume a probability of observing 1 (or H) being p. Then,

- Expectation

$$E(X) = \sum_x xP(x) = 0 \times (1-p) + 1 \times p = p$$

- Variance

$$\begin{aligned} \mathrm{var(X)} &= \sum_x (x - E(X))^2 P(x) = \sum_x (x-p)^2 P(x) \\ &= (0-p)^2 \times (1-p) + (1-p)^2 \times p = p(1-p) \end{aligned}$$

Example: Coin

Let us go back to our coin example where probability of observing 1, or H, is p. Suppose that we perform n independent trials, and are interested in probability that k heads come up in an n-toss sequence, i.e.
$P(k) = P(k \text{ heads come up in a sequence of n tosses})$.

Recall that we have showed already that the probability of any given sequence that contains k heads is $p^k(1-p)^{n-k}$. Hence,

$$P(k) = C(n,k)p^k(1-p)^{n-k}, \quad k = 0, 1, \ldots, n,$$

where $C(n,k) = \frac{n!}{k!(n-k)!}$ and $C(n,k)$ is the number of possible orderings of k successes.

Binomial Distribution

Definition. A r.v. X described as the number of successes in a sequence of n independent Bernoulli trials, follows a **Binomial distribution**, with parameters n (number of trials) and p (probability of success).

Properties. Like in the previous coin example, let us assume a probability of observing 1 (or H) being p and that we have n independent trials. Then,

- Expectation

$$E(X) = E(X_1 + \ldots + X_n) = p + \ldots + p = np$$

- Variance

$$\mathrm{var}(X) = \mathrm{var}(X_1 + \ldots + X_n) = \mathrm{var}(X_1) + \ldots + \mathrm{var}(X_n) = npq$$

Example: Grade of service

Example. An internet service provider has installed c modems to serve the needs of a population of n customers. It is estimated that at a given time, each customer will need a connection with probability p, independently of the others.

What is the probability that there are more customers needing a connection than there are modems?

Example: Grade of service

Solution. Here we are interested in the probability that more than c customers simultaneously need a connection. I.e.,

$$\sum_{k=c+1}^{n} P(k),$$

where

$$P(k) = C(n,k)p^k(1-p)^{n-k}$$

are the binomial probabilities.

This example is typical of problems of sizing the capacity of a facility to serve the needs of a homogeneous population, that consists of independently acting customers. The challenge is to select the size c to achieve a certain threshold probability (sometimes called **grade of service**) that virtually no user is left unserved.

Example: DVD players

A quality control engineer is in charge of testing whether or not 90% of the DVD players produced by his company conform to specifications.

To do this, the engineer randomly selects a batch of 12 DVD players from each day's production. The day's production is acceptable provided no more than 1 DVD player fails to meet specifications. Otherwise, the entire day's production has to be tested.

Q1. What is the probability that the engineer incorrectly passes a day's production as acceptable if only 80% of the day's DVD players actually conform to specification?

Example: DVD players

Solution. Let X denote the number of DVD players in the sample that fail to meet specifications.

Notice that we want to estimate $P(X \leq 1)$ with binomial parameters $n = 12$ and $p = 0.2$.

I.e., we want to find a probability of observing either 0 or 1 defected DVD in a sample of 12 players, while in fact there is 20% of defected DVD in this production series.

$$
\begin{aligned}
P(X \leq 1) &= P(X = 0) + P(X = 1) \\
&= C(12, 0) \times 0.2^0 \times 0.8^{12} + C(12, 01) \times 0.2^1 \times 0.8^{11} \\
&= 0.069 + 0.206 = 0.275
\end{aligned}
$$

Example: DVD players

Q2. What is the probability that the engineer unnecessarily requires the entire day's production to be tested if in fact 90% of the DVD players conform to specifications?

Example: DVD players

Solution. We now want $P(X > 1)$ with parameters $n = 12$ and $p = 0.1$, i.e. that we see more than 1 defected DVD player in a sample of 12 players while in fact 90% of players are good.

$$\begin{aligned} P(X \leq 1) &= P(X = 0) + P(X = 1) \\ &= C(12, 0) \times 0.1^0 \times 0.9^{12} + C(12, 1) \times 0.1^1 \times 0.9^{11} \\ &= 0.659 \end{aligned}$$

Hence,

$$P(X > 1) = 1 - P(X \leq 1) = 0.341.$$

Geometric Distribution

The number of Bernoulli trials needed to get the first cusses has **Geometric distribution**, with parameter p (probability of success in a single trial).

Properties. Like in the previous coin example, let us assume a probability of observing 1 (or H) being p. Then,

- $P(x) = (1-p)^{x-1}p, \quad , x = 1, 2, \ldots$
- Expectation

$$E(X) = \sum_{x=1}^{\infty} x(1-p)^{x-1}p = \frac{p}{(1-q)^2} = \frac{p}{(1-(1-p))^2} = \frac{1}{p}.$$

- Variance

$$\mathrm{var(X)} = \frac{1-\mathrm{p}}{\mathrm{p}^2}.$$

Example: Repeat Until

Example. Examine the following programming statement:

$$\textit{Repeat} \;\; S \;\; \textit{until} \;\; B.$$

What is the probability that S is executed twice?

Solution. Let $P(B = \textit{true}) = 0.1$ and let x be the number of times S is executed. Then, x has a geometric distribution with the probability mass function:

$$P(x) = 0.9^{x-1}0.1.$$

Thus, $P(2) = 0.9 \times 0.1 = 0.09$.

Poisson Distribution

Definition. The number of rare events happening within a fixed period of time has **Poisson distribution**.

Properties. Like in the previous coin example, let us assume a probability of observing 1 (or H) being p. Then,

- $P(x) = e^{-\lambda}\frac{\lambda^x}{x!}, \quad , x = 0, 1, 2, \ldots$
- Expectation
$$E(X) = \lambda.$$
- Variance
$$\mathrm{var}(\mathrm{X}) = \lambda.$$

If a Poisson random variable x has mean λ, we often denote it as $X \sim \mathrm{Poisson}(\lambda)$.

Example: Births

Example. Births in a hospital occur randomly at an average rate of 1.8 births per hour.

What is the probability of observing 4 births in a given hour at the hospital?

Example: Births

Example. Births in a hospital occur randomly at an average rate of 1.8 births per hour.

What is the probability of observing 4 births in a given hour at the hospital?

Solution. Let X be the number of births in a given hour. Notice that the events occur randomly with the mean λ of 1.8. Since we are interested in number of events (i.e. births) happening within a fixed period of time (i.e. hour), $X \sim \mathrm{Poisson}(1.8)$.

Example: Births

Example. Births in a hospital occur randomly at an average rate of 1.8 births per hour.

What is the probability of observing 4 births in a given hour at the hospital?

Solution. Let X be the number of births in a given hour. Notice that the events occur randomly with the mean λ of 1.8. Since we are interested in number of events (i.e. births) happening within a fixed period of time (i.e. hour), $x \sim \mathrm{Poisson}(1.8)$.

We can now use the formula to calculate the probability of observing exactly 4 births in a given hour:

$$P(X = 4) = e^{-1.8} \frac{1.8^4}{4!} = 0.0732.$$

Example: Births – CONTD

Example. What about the probability of observing more than or equal to 2 births in a given hour at the hospital?

Example: Births – CONTD

Example. What about the probability of observing more than or equal to 2 births in a given hour at the hospital?

Solution. We want to find:

$$P(X \geq 2) = P(X = 2) + P(X = 3) + \ldots,$$

i.e. an infinite number of probabilities to calculate. Can we do it?

Example: Births – CONTD

At the same notice that

$$
\begin{aligned}
P(X \geq 2) &= P(X = 2) + P(X = 3) + \ldots \\
&= 1 - P(X < 2)
\end{aligned}
$$

Example: Births – CONTD

At the same notice that

$$
\begin{aligned}
P(X \geq 2) &= P(X = 2) + P(X = 3) + \dots \\
&= 1 - P(X < 2) \\
&= 1 - \Big(P(X = 0) + P(X = 1)\Big) \\
&= 1 - \left(e^{-1.8}\frac{1.8^0}{0!} + e^{-1.8}\frac{1.8^1}{1!}\right) \\
&= 1 - (0.16529 + 0.29753) \\
&= 0.537.
\end{aligned}
$$