Discrete random variables

Chrysafis Vogiatzis

Lectures 5 and 6

Learning objectives

After these lectures, we will be able to:

- Define discrete and continuous random variables.
- Differentiate between discrete and continuous random variables.
- Differentiate between when to use cumulative distribution functions and probability mass functions.
- Give examples of at least four different discrete distributions.
- Recall when to and how to use:
  - binomially distributed random variables.
  - geometric distributed random variables.
  - hypergeometric distributed random variables.
  - Poisson distributed random variables.

Motivation: 2-engine vs. 4-engine aircraft

Suppose that for a flight to be completed successfully (which we would really *love*) we need at least half of the engines to be operational at the end of the trip. In the case of a 2-engine aircraft, this means at least one; in a 4-engine aircraft, we'd need at least two. We are ordering engines from a production company and we would like to see whether buying two (for a 2-engine plane) or four (for a 4-engine plane). Which one would be safer?

Motivation: Big in Japan

Over the last 135 years, there have been 5 earthquakes of seismic intensity over 7.0 in the Kanto region of Japan. However, especially for those of us living in seismogenic zones [1], we probably grew up hearing statements such as "the probability of an earthquake in X within the next Y years is $Z\%$". Hey, this is what we do in this class!

[1] An area with high seismic/earthquake activity.

We make the following assumptions for big earthquakes:

1. Big earthquakes are independent events – that is the fact that a big

earthquake happened does not increase or decrease the probability of another big earthquake soon.

2. The probability of an earthquake occurring is the same throughout the year [2].

[2] This is a property also called homogeneity. More on that later.

What is the probability that there will be one big earthquake in the Kanto region in the next year? What is the probability that there will be one big earthquake in the Kanto region in the next decade?

Random variables

The world around us is a series of random processes, whose outcomes affect the way we perceive things. Mathematically, we need to somehow define these outcomes – using numerical representations.

Definition 1 (Random variables) *With the term random variable [3] we mean a* ***real-valued function*** *defined over the sample space.*

[3] Also termed random quantities or stochastic variables.

Definition 2 (Random variables) *A random variable is a function that associates a number with each element of the sample space.*

Classification

In the next few lectures, we will separate our discussion between discrete and continuous random variables.

- Discrete random variables take countable, discrete values. [4]
- Continuous random variables can take any real-value. [5]

[4] For example, the side of a die, the number of customers.

[5] For example, the time until the next bus arrives, the lifetime of a light bulb, the pressure of a gas.

Recall that we had made that distinction in the past for random experiments and their sample spaces! [6]

[6] See Lecture 1.

> Classify these random variables as discrete or continuous:
>
> 1. the time it takes for a biker to go from one side of campus to the other.
> 2. the number of red lights the biker has to stop at when going from one side of the campus to the other.
> 3. the distance a biker traverses to go from one side of campus to the other.
> 4. the number of times the biker changed speed gear while going from one side of campus to the other.

Functions

When a random variable behaves a specific way, we say that it follows a **probability distribution**. A probability distribution is typically described by two distribution functions:

- The **probability mass function** for discrete random variables or **probability density function** for continuous random variables.
- The **cumulative distribution function**.

We formally define those where needed for discrete and continuous random variables.

The remainder for our lecture notes is devoted to discrete random variables. See Lectures 7 and 8 for a discussion on continuous random variables.

Discrete random variables

Let X be a **discrete** random variable.

Definition 3 *We define the* **probability mass function** *(pmf)* $p(x)$ [7] *of a discrete random variable* X *as the probability that it takes a specific value* x*:*

$$p(x) = P(X = x).$$

[7] Some textbooks may use $f(x)$ for the probability mass function. In our notes, we will use $p(x)$ for discrete random variables and their probability mass functions.

One thing we need to be very careful with:

- Distinguish between X (upper case X) and x (lower case x)! X is the random variables; x is a given value. [8]

[8] Example: we can write $P(X = 7)$ and read "what is the probability that *random variable* X is equal to the *value* 7"?

For example, the question "what is the probability that a store has **exactly** 20 customers enter in the next hour?" can be addressed using the **probability mass function** as follows. First, let X be a random variable that represents the number of customers that enter the store in the next hour. Then, express the probability as $P(X = 20)$.

Definition 4 *We define the* **cumulative distribution function** *(cdf)* $F(x)$ *of a discrete random variable as the probability that it takes* up to *a value* x*, i.e.,*

$$F(x) = P(X \leq x) = \sum_{y:y \leq x} P(X = y) = \sum_{y:y \leq x} p(y).$$

For example, the question "what is the probability that a store has **up to** 20 customers enter in the next hour?" can be addressed using the **cumulative distribution function** as follows. First, let X be a random variable that represents the number of customers that enter the store in the next hour. Then, express the probability as

$$
\begin{aligned}
F(20) = P(X \leq 20) = & \\
&= P(X=0) + P(X=1) + \ldots + P(X=20) = \\
&= \sum_{y \leq 20} P(X=y).
\end{aligned}
$$

An immediate result from the definition of the cdf is that if we are interested in the probability of seeing more than a certain value x we may write:

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

Combining the two definitions (of $P(X \leq x)$ and $P(X > x)$) we get that the probability of X taking values between a and b is [9]:

$$P(a \leq X \leq b) = F(b) - F(a).$$

[9] This is easy to work out. It is left as an exercise to the reader.

Defining these two functions helps us classify random variables based on their properties, as we will spend the rest of the lectures finding out.

2-engine vs. 4-engine

We saw that a plane performs a trip safely if at least half of its engines are operational. Let X be the number of engines that have failed during a trip. Then, we should be looking for:

- $P(X \leq 1)$: for the probability of a successful trip with a 2-engine aircraft.
- $P(X \leq 2)$: for the probability of a successful trip with a 4-engine aircraft.

Earthquake probabilities

In a similar fashion, let X be the number of earthquakes in the Kanto region in the next year and Y be the same number in the next decade. Then, we should be looking for:

- $P(X = 1)$: for the probability of one big earthquake in the next year.
- $P(Y = 1)$: for the probability of one big earthquake in the next decade.

Should you use the pmf or the cdf?

- To avoid paying your friends in a game of Monopoly you need to get a 6 or less when throwing two dies.
- An exam has 10 multiple choice questions. What is the probability you answer all of them correctly?
- An exam has 10 multiple choice questions. What is the probability you answer more than or equal to 8 questions correctly?
- Two people are playing a game that is best out of three. What is the probability the first player wins with a score of 2 to 1?

Assume a discrete random variable X with n outcomes x_i, $i = 1, \ldots, n$. Then, the probability mass function $p(x)$ of random variable X has to satisfy the following three rules:

1. $p(x_i) = P(X = x_i)$, for every outcome $x_i, i = 1, \ldots, n$.
2. $p(x_i) \geq 0$.
3. $\sum\limits_{i=1}^{n} p(x_i) = 1$

Urns and balls

Two balls are drawn from an urn containing 5 red and 4 black balls. Define a random variable X as the number of red balls drawn. What is its probability mass function?

- X has three outcomes: $0, 1, 2$.
- To select 0 red balls: $p(0) = P(X = 0) = \frac{C_{4,2}}{C_{9,2}} = \frac{6}{36}$
- To select 1 red ball: $p(1) = P(X = 1) = \frac{C_{4,1} \cdot C_{5,1}}{C_{9,2}} = \frac{20}{36}$
- Finally, for 2 red balls: $p(2) = P(X = 2) = \frac{C_{5,2}}{C_{9,2}} = \frac{10}{36}$

We may verify that these probabilities satisfy all three rules of a valid probability mass function.

Urns and balls

Two balls are drawn from an urn containing 5 red and 4 black balls. Define a random variable X as the number of red balls drawn. What is its probability mass function?

- X has three outcomes: $0, 1, 2$.
- To select 0 red balls: $p(0) = P(X = 0) = \frac{C_{4,2}}{C_{9,2}} = \frac{6}{36}$
- To select 1 red ball: $p(1) = P(X = 1) = \frac{C_{4,1} \cdot C_{5,1}}{C_{9,2}} = \frac{20}{36}$
- Finally, for 2 red balls: $p(2) = P(X = 2) = \frac{C_{5,2}}{C_{9,2}} = \frac{10}{36}$

We may verify that these probabilities satisfy all three rules of a valid probability mass function.

Calculating probabilities

A sample space is described by two mutually exclusive outcomes A and B. We have observed that for some real number x, the pmf is $P(A) = 3 \cdot x$ and $P(B) = 10 \cdot x^2$. What is x?
This type of question needs us to use the pmf rules. We need to verify what x should be in order to satisfy $P(A), P(B) \geq 0$ and $P(A) + P(B) = 1$. Replacing the pmf in the second equality we have:

$$\begin{aligned} P(A) + P(B) = 1 &\implies 3 \cdot x + 10 \cdot x^2 = 1 \\ &\implies 10 \cdot x^2 + 3 \cdot x - 1 = 0 \implies \\ &\implies x = \begin{cases} \text{-0.5} \\ \\ 0.2 \end{cases} \end{aligned}$$

Replacing $x = -0.5$, we get that $P(B) = 2.5$, but $P(A) = -1.5 < 0$. Replacing $x = 0.2$, we get that $P(A) = 0.6$ and $P(A) = 0.4$, and is the correct answer.

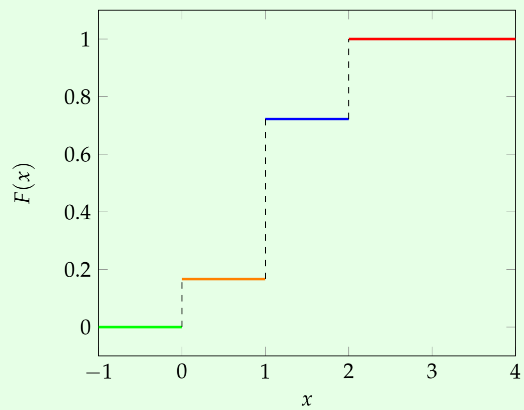
The cumulative distribution function (cdf) of a discrete random variable X needs to satisfy in turn two rules:

1. $0 \leq F(x) \leq 1$.
2. If $x \leq y$, then $F(x) \leq F(y)$.

Urns and balls

Consider the previous sample space $0, 1, 2$ with $p(0) = \frac{6}{36}, p(1) = \frac{20}{36}, p(2) = \frac{10}{36}$. Then:

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{6}{36}, & 0 \leq x < 1 \\ \frac{26}{36}, & 1 \leq x < 2 \\ 1, & x \geq 2. \end{cases}$$



Let the sample space be $S = \{1, 2, 3\}$ with $p(1) = \frac{1}{2}, p(2) = \frac{1}{3}, p(3) = \frac{1}{6}$.

- Verify this is a valid pmf.
- Write the cdf $F(x)$.
- Draw the cdf (like we showed in the previous example).

The binomial distribution

Before we get to the **binomial distribution**, we need to introduce **Bernoulli random variables**. This first random variable we will introduce is also (probably) the simplest!

Consider a single experiment that has only two probable outcomes:

1. **success** which happens with probability p; and
2. **failure** which happens with probability $q = 1 - p$.

Now, define a random variable X based on that single experiment:

$$X = \begin{cases} 0, & \text{if the experiment failed;} \\ 1, & \text{if the experiment succeeded.} \end{cases}$$

The key is that we consider only **one** experiment [10]. For the Bernoulli distribution, we have:

$$\text{pmf:} \quad \begin{array}{l} P(X=0) = q = 1-p \\ P(X=1) = p \end{array}$$

$$\text{cdf:} \;\; F(x) = \begin{cases} 0, & x < 0 \\ 1-p, & x < 1 \\ 1, & x \geq 1 \end{cases}$$

[10] Will the next coin toss be a heads (success) or a tail (failure)? Will it rain (success) or not (failure)? Will my favorite NBA team win (success) its next game or not (failure)? Will the next patient be cured (success) or not (failure)?

Urns and balls

An urn contains 40 black and 10 red balls. You pick at random one ball from the urn. Let X be the number of black balls you pick from the urn. What is the pmf of X?

X is a Bernoulli distributed random variable with pmf:

- $P(X=0) = \frac{10}{50} = 0.2$.
- $P(X=1) = \frac{40}{50} = 0.8$.

What if we consider more than one experiments? What if, say, we picked 5 balls and wanted to get 3 black ones [11] This is where **binomially distributed random variables** come in play! The setup is simple:

[11] Multiple experiments could mean multiple coin tosses, or a control group of 100 patients, or a best-of-five game series!

- n independent experiments/trials.
- each experiments ends up in a success with probability p and a failure with probability $1-p$;
  - that is, each trial is a Bernoulli random variable.
- Let X be the number of successes.

Then X is a binomial random variable. We may also write that $X = binom(n, p)$ as n (number of experiments) and p (probability of success in each individual experiment) are the only necessary parameters to fully define this random variable.

Coin tosses

The most common example to explain binomial random variables comes from coin tosses. Assume we possess a "fair" coin with probability of Heads $p = 0.5$, and probability of Tails $q = 1-p = 0.5$? What is the probability that there will be exactly 2 Heads in $n = 3$ tosses of the coin? This would be a binomial distribution with $n = 3$, $p = 0.5, q = 0.5$, and $x = 2$ Heads.

The formula for calculating the probability for binomially distributed random variables is: [12]

$$\boxed{p(x) = P\left(X = x\right) = \tbinom{n}{x} \cdot p^x \cdot (1-p)^{n-x}, \quad \text{for } x = 0, 1, \ldots, n.}$$

Recall that $\binom{n}{x} = \frac{n!}{x!\cdot(n-x)!}$ as we had seen in a previous lecture. [13]

[12] The derivation of the formula is part of Lecture 5's worksheet.

[13] See Lecture 2.

Coin tosses

We may now address the earlier question:

$$p(2) = \binom{3}{2} \cdot 0.5^2 \cdot (1-0.5)^{3-2} = \frac{3!}{2!\cdot 1!} \cdot 0.25 \cdot 0.5 = 0.375.$$

2-engine vs. 4-engine

In our motivational example, let p be the probability that an engine fails during a trip, and hence $q = 1 - p$ is the probability it does not fail. For the success of each plane, we have:

2-engine:

$$\begin{aligned} P(X \le 1) &= P(X=0) + P(X=1) = \\ &= \binom{2}{0} \cdot p^0 \cdot (1-p)^2 + \binom{2}{1} \cdot p^1 \cdot (1-p)^1 = \\ &= 1 - 2p + p^2 + 2 \cdot p - 2 \cdot p^2 = 1 - p^2 \end{aligned}$$

4-engine:

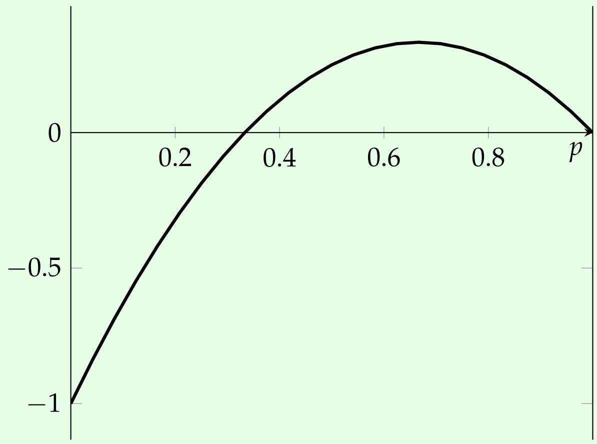
$$\begin{aligned} &P(X \le 2) = P(X=0) + P(X=1) + P(X=2) = \\ &= \binom{4}{0} \cdot p^0 \cdot (1-p)^4 + \binom{4}{1} \cdot p^1 \cdot (1-p)^3 + \binom{4}{2} \cdot p^2 \cdot (1-p)^2 \\ &= 1 - 4 \cdot p + 6 \cdot p^2 - 4 \cdot p^3 + p^4 + 4 \cdot p - 12 \cdot p^2 + 12 \cdot p^3 - 4 \cdot p^4 \\ &+ 6 \cdot p^2 - 12 \cdot p^3 + 6 \cdot p^4 = 1 + 3 \cdot p^4 - 4 \cdot p^3 \end{aligned}$$

Can we compare the two? To prefer a 2-engine plane we need its probability of success to be higher, that is we need:

$$\begin{aligned} 1 - p^2 \ge 1 + 3 \cdot p^4 - 4 \cdot p^3 \implies -3 \cdot p^4 + 4 \cdot p^3 - p^2 \ge 0 \implies \\ \implies -3 \cdot p^2 + 4 \cdot p - 1 \ge 0. \end{aligned}$$

Let's plot $y = -3 \cdot p^2 + 4 \cdot p - 1$ and see what we get! Since for $y \ge 0$ we prefer a 2-engine aircraft, it suffices to see when y is nonnegative in the plot!

2-engine vs. 4-engine (cont'd)



We observe that for probability of engine failure $p \geq \frac{1}{3}$, then a 2-engine plane is favored!

An urn contains 40 black and 10 red balls. You pick at random one ball from the urn, check its color, and after checking its color, you put it back in the urn. Let X be the number of black balls you pick from the urn in $n = 10$ tries. What is the probability that $X = 6$? What is the probability that $X \geq 9$?

Food for thought: why was it important in the previous example to put the ball back in the urn? What changes if I remove it from the urn?

The geometric distribution

Let's look at another extension of Bernoulli random variables. Earlier, during our discussion for binomially distributed random variables, we cared about the number of successes in a series of trials. How about the **first success** though? When did it occur? Since we are talking about a series of experiments, this first success can occur at the first, second, third, and so on, try.

Coin tosses

Assume again we are in possession of a fair coin. What is the probability the first Heads appears after three tries?

In general, we have that the probability that the first success is seen after exactly x trials is: [14]

$$P(X = x) = (1 - p)^{x-1} \cdot p.$$

[14] The derivation of this formula is also, albeit easier, part of Lecture 5's worksheet.

Learning basketball

A kid learning basketball is shooting free throws with a probability of scoring equal to 25%. What is the probability the kid has to shoot four free throws until scoring the first one?

The number of free throws until the first one is scored X is a geometric random variable with $p = 0.25$ and $x = 4$, hence we have

$$P(X = 4) = (1 - 0.25)^3 \cdot 0.25 = 0.75^3 \cdot 0.25 = 0.1055.$$

Assume we have a fair coin. What is the probability..

- the first Heads appears in the 2nd toss?
- the first Heads appears in the 5th toss?
- the first Heads appears in the 10th toss?

The hypergeometric distribution

What if..

- we had N items;
- $K \leq N$ of them are successes (the remaining $N - K$ are failures);
- we drew n of them;
- what is the probability we get k successes in the sample of n?

We have actually dealt with this problem before. Recall the example from Lecture 2:

Quality control

A package is set to leave a factory and be sent to a retailer. The package contains 100 items. We already know that exactly 3 of the 100 items are defective. The quality control team over at the retailer works as follows: they select a sample of 6 items from the 100, and check them. If there are 0 defective items in the selected sample of 6, they accept the package and sell its contents; otherwise, they send the package back. What is the probability that the quality control rejects the package and sends it back?

The answer to this probability was $0.171 = 17.1\%$. Now, check how this problem matches the setup of the hypergeometric distribution: $N = 100$ total population size; $K = 3$ defective ones; $n = 6$ sample size. If X is the number of defective items picked in the sample, then the pmf for the hypergeometric is:

$$\text{pmf:} \;\; P(X = x) = \frac{\binom{K}{x} \cdot \binom{N-K}{n-x}}{\binom{N}{n}}$$

> An urn contains 40 black and 10 red balls. You pick at random a sample of five balls from the urn. Let X be the number of black balls in the sample. What is the probability that $X = 3$?

The big difference between the binomial and the hypergeometric distribution is in the **sampling with replacement** [15] and the **sampling without replacement** [16] The main difference is that with replacement, the probability of picking an item stays the same throughout the experiment, no matter how many times it is repeated; without replacement, the probability changes with every selection.

[15] For example, taking 5 balls from the urn one-by-one, looking at each one's color, and putting it back in, before picking the next; or selecting 5 items from a box one-by-one, checking it, and placing it back in again before picking the next.

[16] For example, taking 5 balls from the urn at the same time and looking at their colors together; or selecting 5 items from a box at the same time, checking them and seeing if they are defective.

The Poisson distribution

We do the opposite of what we normally do: we will motivate the Poisson distribution from a mathematical perspective instead of through an example. Recall the binomial distribution and its probability mass function:

$$p(x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}\,.$$

Assume that we try way too many experiments (in essence let $n \to \infty$) and define λ as the number of successes we get: that is, $p = \frac{\lambda}{n}$. Let's replace this in the probability mass function itself:

$$\begin{aligned} p(x) &= \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} \\ &= \frac{n!}{x! \cdot (n-x)!} \cdot \left(\frac{\lambda}{n}\right)^x \cdot \left(1 - \frac{\lambda}{n}\right)^{n-x}\,. \end{aligned}$$

Let us now employ some of the cool limit properties that we know as $n \to \infty$!

$$\begin{aligned}
\lim_{n\to\infty} P(X = x) &= \lim_{n\to\infty} \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} = \\
&= \lim_{n\to\infty} \frac{n!}{x! \cdot (n-x)!} \cdot \left(\frac{\lambda}{n}\right)^x \cdot \left(1 - \frac{\lambda}{n}\right)^{n-x} = \\
&= \lim_{n\to\infty} \frac{\lambda^x}{x!} \cdot \underbrace{\frac{n!}{(n-x)! \cdot n^x}}_{1} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{e^{-\lambda}} \cdot \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-x}}_{1} = \\
&= e^{-\lambda} \frac{\lambda^x}{x!}
\end{aligned}$$

Definition 5 *A discrete random variable X taking values $0, 1, 2, \dots$ is a Poisson random variable with parameter (rate) $\lambda > 0$ if:*

$$p(x) = P\left(X = x\right) = e^{-\lambda} \frac{\lambda^x}{x!}.$$

Poisson random variables have a wide, *wide* array of applications. They have been used to model:

- the number of phone calls that a call center gets every day.
- the number of shark attacks in California every year.
- the number of home runs in a baseball series.
- the number of patients arriving in an emergency department every night.
- the number of website requests per second.
- the number of earthquakes expected to hit a seismogenic area every decade.

As can be seen from the examples, Poisson distributed random variables are commonly used to model the number of events that happen in a given interval. Poisson distributed random variables need to satisfy three main conditions:

1. independence: an event happening should not affect the rate with which more events happen.
2. homogeneity: the rate with which events happen is constant.
3. no two events can occur at *exactly* the same time. Instead there is a small interval of time that separates two consecutive events.

Earthquake probabilities

We will model our motivating example with predicting the probability of an earthquake in the Kanto region of Japan using the Poisson distribution. We need to estimate λ, the rate of events. From the data, we are told that there have been 5 big earthquakes over the last 135 years, and hence:

$$\lambda = \frac{5}{135} = 0.037 \text{ earthquakes per year.}$$

We are interested in:

- $P(X = 1)$: for the probability of one big earthquake in the next year.

$$P(X = 1) = e^{-0.037} \cdot \frac{0.037^1}{1!} = 0.0357 = 3.57\%.$$

When interested in the probability of one big earthquake over the next decade:

- $P(Y = 1)$: for the probability of one big earthquake in the next decade.

Here we need to adapt the rate to accommodate periods of 10 years. Hence, $\lambda = 0.37$ earthquakes per year. Finally:

$$P(Y = 1) = e^{-0.37} \cdot \frac{0.37^1}{1!} = 0.2556 = 25.56\%.$$

Finally, let's address the probability that there is *at least one* big earthquake in the next decade:

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - e^{-0.37} \cdot \frac{0.37^0}{0!} =$$
$$= 1 - 0.6907 = 0.3093 = 30.93\%.$$

Is it fair to assume that typos appearing in notes follow a Poisson distribution? Why/Why not?

Assume typos appear in my notes following a Poisson distribution with a rate of $\lambda = 0.5/page$. What is the probability that no typos exist in the first page? What is the probability that there exist more than 1 typo in the first 10 pages?

Plotting the Poisson distribution also proves an interesting endeavor. Let's remember that all distributions we are discussing are discrete: hence, we will simply plot each point and then connect the points with a line. For example, in Figure 1, we show the case for $\lambda = 1$ and how we would connect the different data points.

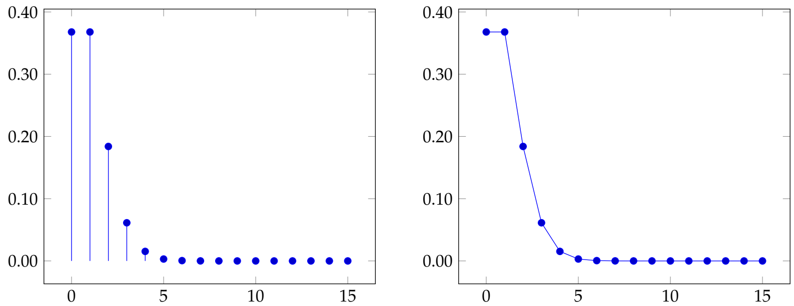


Figure 1: The Poisson distribution for $\lambda = 1$.

We do the same for $\lambda = 2, 5, 10$ in Figures 2, 3, 4.

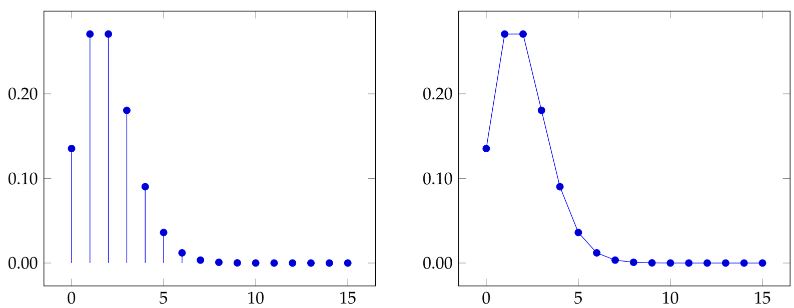


Figure 2: The Poisson distribution for $\lambda = 2$.

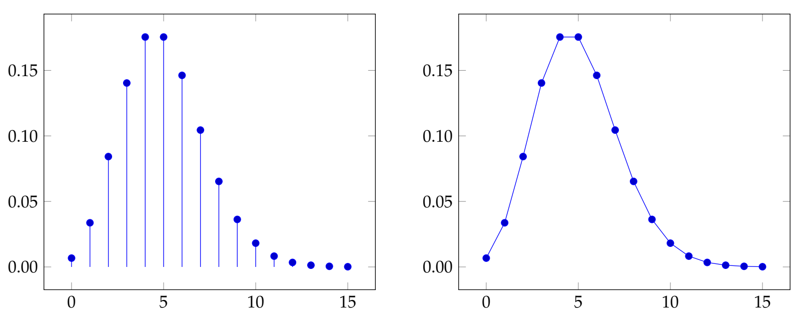


Figure 3: The Poisson distribution for $\lambda = 5$.

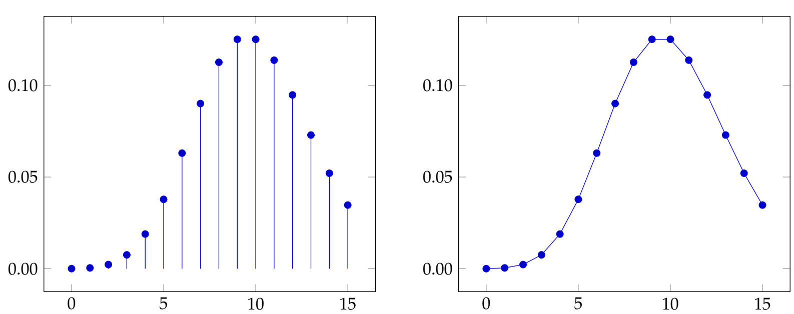


Figure 4: The Poisson distribution for $\lambda = 10$.

Finally, take a look at Figure 5. See what happens when λ takes on very big values...

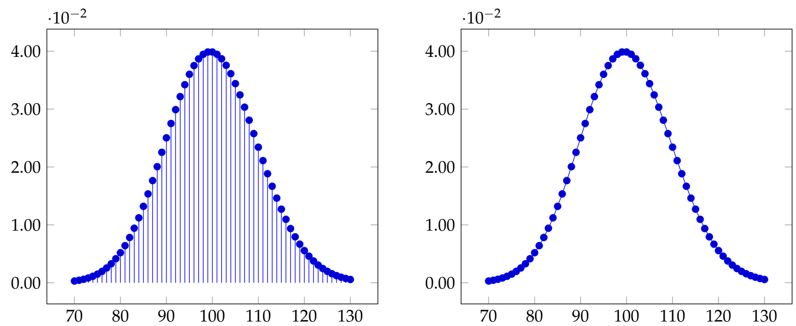


Figure 5: The Poisson distribution for $\lambda = 100$.

The uniform distribution

Finally, we see the simplest discrete distribution, the uniform distribution. Think of a discrete random variable with n different outcomes $x_i, i = 1, \ldots, n$. Now assume that:

- all n outcomes are equally probable, then we have a uniform random variable.
- each of the outcomes is equally probable, i.e., $p_i = P(X = x_i) = \frac{1}{n}$.

In a special case, the discrete random variable take *integer* values in $[a, b]$. In that case, the pmf is

$$p_i = \frac{1}{b - a + 1}, \text{ for all } i = a, a + 1, \ldots, b.$$

A diamond cutting facility

A demanding customer has shown up in a diamond cutting facility and has asked for 2 *custom-made fine-cut diamond castings*. They are willing to buy 2 of those, as long as they are of high quality.. Diamond cutting is an expensive process, but you can make a lot of money out of it, and hence decide to take on the order. You plan to buy enough material for $Q = 4$ castings, just to be safe. Assuming that diamond cutting is a purely random process and all outcomes (producing $x = 0, 1, \ldots, Q$ high quality diamonds) are equally probable. What is the probability you satisfy your customer?

We need $x \geq 2$ high-quality fine-cut castings. The number of high-quality castings produced follows a uniform distribution, so:

$$P(X = x) = \frac{1}{Q + 1} = \frac{1}{5}.$$

Hence, to satisfy the customer we have a probability of:

$$P(X \geq 2) = P(X = 2) + P(X = 3) + P(X = 4) = \frac{3}{5}.$$

Summary

In Table 1, we provide all of our results from Lectures 5 and 6. One could simply refer to these (and the keyword at the end of the page) for all information about discrete probability distributions.

Table 1: A summary of all results from Lectures 5 and 6.

Name	**Parameters**	**Values**	**pmf**
Bernoulli	$0 < p < 1$	$\{0, 1\}$	$p(0) = 1 - p$ $p(1) = p$
Binomial	$0 < p < 1$, $n \geq 0$	$\{0, 1, \ldots, n\}$	$p(x) = \binom{n}{x} p^x \cdot (1-p)^{n-x}$
Geometric	$0 < p < 1$	$\{1, 2, \ldots\}$	$p(x) = (1-p)^{x-1} \cdot p$
Hypergeometric	$N, K, n \geq 0$	$\{1, 2, \ldots\}$	$p(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$
Poisson	$\lambda > 0$	$\{0, 1, \ldots\}$	$p(x) = e^{-\lambda} \frac{\lambda^x}{x!}$
Uniform	-	$[a, b]$	$p(x) = \frac{1}{b - a + 1}$

Some keywords that might help you narrow down your search.

Bernoulli: "one single experiment/trial"; "success/failure"; "p and $q = 1 - p$".

Binomial: "multiple experiments/trials"; "success/failure"; "probabilities stay the same from experiment to experiment"; "how many successes in n tries?"; "with replacement".

Geometric: "number of experiments/trials until first success"; "success/failure"; "probabilities stay the same from experiment to experiment".

Hypergeometric: "sample"; "success/failure"; "without replacement"; "how many successes in a sample of size n?".

Poisson: "rate of events"; "number of events in an interval".

Uniform: "equally probable"; "outcomes are integer in $[a, b]$".