CS 107, Probability, Spring 2020
Lecture 24

Michael Poghosyan
mpoghosyan@aua.am

AUA

27 March 2020

Content

- Some important Discrete rv.s
  - Poisson Distribution;
  - Negative Binomial Distribution;
  - Hypergeometric Distribution

Poisson Distribution

Poisson Distribution

Distribution Name: *Pois*, *Poisson*;
Parameters: λ $(\lambda > 0)$

Poisson Distribution

We say that the r.v. X has a Poisson Distribution with the parameter (rate per unit of measure) λ $(\lambda > 0)$, and we will write $X \sim Pois(\lambda)$, if the PMF of X is given by

$$\mathbb{P}(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \dots$$

that is,

Values of X	0	1	2	$\dots$	k	$\dots$
$\mathbb{P}(X = x)$	$e^{-\lambda}$	$e^{-\lambda} \cdot \frac{\lambda^1}{1!}$	$e^{-\lambda} \cdot \frac{\lambda^2}{2!}$	$\dots$	$e^{-\lambda} \cdot \frac{\lambda^k}{k!}$	$\dots$

Poisson Distribution

Distribution Name: *Pois*, *Poisson*;
Parameters: λ $(\lambda > 0)$

Poisson Distribution

We say that the r.v. X has a Poisson Distribution with the parameter (rate per unit of measure) λ $(\lambda > 0)$, and we will write $X \sim Pois(\lambda)$, if the PMF of X is given by

$$\mathbb{P}(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \ldots$$

that is,

Values of X	0	1	2	$\ldots$	k	$\ldots$
$\mathbb{P}(X = x)$	$e^{-\lambda}$	$e^{-\lambda} \cdot \frac{\lambda^1}{1!}$	$e^{-\lambda} \cdot \frac{\lambda^2}{2!}$	$\ldots$	$e^{-\lambda} \cdot \frac{\lambda^k}{k!}$	$\ldots$

Exercise: Use $Calc2 \cup RA$ to prove that $\sum_{k=0}^{\infty} \mathbb{P}(X = k) = 1$.

Poisson Distribution: Interpretation

Usage: Poisson Distribution is modeling the number of times a rare event occurs in a fixed interval of time or fixed area of the space, and λ is the rate of that event per unit of a measure (say, per minute, per hour, per m^2, cm^2 etc.).

Poisson Distribution: Interpretation

Usage: Poisson Distribution is modeling the number of times a rare event occurs in a fixed interval of time or fixed area of the space, and λ is the rate of that event per unit of a measure (say, per minute, per hour, per m^2, cm^2 etc.).

Notes:

- λ, the rate parameter, is the average number of events per unit of a measure (per unit of time, per unit of area);

Poisson Distribution: Interpretation

Usage: Poisson Distribution is modeling the number of times a rare event occurs in a fixed interval of time or fixed area of the space, and λ is the rate of that event per unit of a measure (say, per minute, per hour, per m^2, cm^2 etc.).

Notes:

- λ, the rate parameter, is the average number of events per unit of a measure (per unit of time, per unit of area);
- If X is a Poisson r.v., then it's range is $\{0, 1, 2, 3, 4, ...\}$;

Poisson Distribution: Interpretation

Usage: Poisson Distribution is modeling the number of times a rare event occurs in a fixed interval of time or fixed area of the space, and λ is the rate of that event per unit of a measure (say, per minute, per hour, per m^2, cm^2 etc.).

Notes:

- λ, the rate parameter, is the average number of events per unit of a measure (per unit of time, per unit of area);
- If X is a Poisson r.v., then it's range is $\{0, 1, 2, 3, 4, ...\}$;
- Do not read Poisson as **Poison** ! $\ddot\smile$

Poisson Distribution: Some Examples

For example, we can model the following events by using the Poisson distribution:

Poisson Distribution: Some Examples

For example, we can model the following events by using the Poisson distribution:

- The number of emails received in a day

Poisson Distribution: Some Examples

For example, we can model the following events by using the Poisson distribution:

- The number of emails received in a day
- The number of phone calls in a day for a person (or some call center)

Poisson Distribution: Some Examples

For example, we can model the following events by using the Poisson distribution:

- The number of emails received in a day
- The number of phone calls in a day for a person (or some call center)
- The number of car accidents at some place between 9AM and 9PM

Poisson Distribution: Some Examples

For example, we can model the following events by using the Poisson distribution:

- The number of emails received in a day
- The number of phone calls in a day for a person (or some call center)
- The number of car accidents at some place between 9AM and 9PM
- The number of misprints in a page of a published Probability book (or in a newspaper);

Poisson Distribution: Some Examples

For example, we can model the following events by using the Poisson distribution:

- The number of emails received in a day
- The number of phone calls in a day for a person (or some call center)
- The number of car accidents at some place between 9AM and 9PM
- The number of misprints in a page of a published Probability book (or in a newspaper);
- The number or defective items in a production batch;

Poisson Distribution: Some Examples

For example, we can model the following events by using the Poisson distribution:

- The number of emails received in a day
- The number of phone calls in a day for a person (or some call center)
- The number of car accidents at some place between 9AM and 9PM
- The number of misprints in a page of a published Probability book (or in a newspaper);
- The number or defective items in a production batch;
- The number of network failures in a day;

Poisson Distribution: Some Examples

For example, we can model the following events by using the Poisson distribution:

- The number of emails received in a day
- The number of phone calls in a day for a person (or some call center)
- The number of car accidents at some place between 9AM and 9PM
- The number of misprints in a page of a published Probability book (or in a newspaper);
- The number or defective items in a production batch;
- The number of network failures in a day;
- The number of earthquakes in a year;

Poisson Distribution: Some Examples

For example, we can model the following events by using the Poisson distribution:

- The number of emails received in a day
- The number of phone calls in a day for a person (or some call center)
- The number of car accidents at some place between 9AM and 9PM
- The number of misprints in a page of a published Probability book (or in a newspaper);
- The number or defective items in a production batch;
- The number of network failures in a day;
- The number of earthquakes in a year;
- The number of goals in a football game;

Poisson Distribution: Some Examples

For example, we can model the following events by using the Poisson distribution:

- The number of emails received in a day
- The number of phone calls in a day for a person (or some call center)
- The number of car accidents at some place between 9AM and 9PM
- The number of misprints in a page of a published Probability book (or in a newspaper);
- The number or defective items in a production batch;
- The number of network failures in a day;
- The number of earthquakes in a year;
- The number of goals in a football game;
- etc.

Poisson Distribution: Note

When is the Poisson distribution an appropriate model?

Poisson Distribution: Note

When is the Poisson distribution an appropriate model? *Wiki* gives the following "test".

Poisson Distribution: Note

When is the Poisson distribution an appropriate model? *Wiki* gives the following "test". The Poisson distribution is an appropriate model if the following assumptions are true

Poisson Distribution: Note

When is the Poisson distribution an appropriate model? *Wiki* gives the following "test". The Poisson distribution is an appropriate model if the following assumptions are true

- There is a fixed interval of time or space during or in which we are considering our events (say, every minute, every day, every month, every day from 10:30AM till 11:20AM, every square cm etc.)

Poisson Distribution: Note

When is the Poisson distribution an appropriate model? *Wiki* gives the following "test". The Poisson distribution is an appropriate model if the following assumptions are true

- There is a fixed interval of time or space during or in which we are considering our events (say, every minute, every day, every month, every day from 10:30AM till 11:20AM, every square cm etc.)
- X is the number of times an event occurs in an interval and X can take values 0, 1, 2, ...

Poisson Distribution: Note

When is the Poisson distribution an appropriate model? *Wiki* gives the following "test". The Poisson distribution is an appropriate model if the following assumptions are true

- There is a fixed interval of time or space during or in which we are considering our events (say, every minute, every day, every month, every day from 10:30AM till 11:20AM, every square cm etc.)
- X is the number of times an event occurs in an interval and X can take values 0, 1, 2, ...
- The occurrence of one event does not affect the probability that a second event will occur. That is, events occur independently.

Poisson Distribution: Note

When is the Poisson distribution an appropriate model? *Wiki* gives the following "test". The Poisson distribution is an appropriate model if the following assumptions are true

- There is a fixed interval of time or space during or in which we are considering our events (say, every minute, every day, every month, every day from 10:30AM till 11:20AM, every square cm etc.)
- X is the number of times an event occurs in an interval and X can take values 0, 1, 2, ...
- The occurrence of one event does not affect the probability that a second event will occur. That is, events occur independently.
- The rate at which events occur is constant. The rate cannot be higher in some intervals and lower in other intervals.

Poisson Distribution: Note

When is the Poisson distribution an appropriate model? *Wiki* gives the following "test". The Poisson distribution is an appropriate model if the following assumptions are true

- There is a fixed interval of time or space during or in which we are considering our events (say, every minute, every day, every month, every day from 10:30AM till 11:20AM, every square cm etc.)
- X is the number of times an event occurs in an interval and X can take values 0, 1, 2, ...
- The occurrence of one event does not affect the probability that a second event will occur. That is, events occur independently.
- The rate at which events occur is constant. The rate cannot be higher in some intervals and lower in other intervals.
- Two events cannot occur at exactly the same instant.

Poisson Distribution: Examples

Example 24.1: Which of the following r.v.s can be modeled as a Poisson r.v.?

a. $X =$ number of cars passing by my window in an hour;
b. $Y =$ number of cars passing by my window between 10AM and 11AM;
c. $Z =$ no of calls at taxi center per hour;
d. $T =$ no of call at some emergency center per hour;
E. $U =$ daily number of absent students from Probability class.

Poisson Distribution: Examples

Example 24.2: Assume

$$X \sim Pois(4)$$

Example 24.2: Assume

$$X \sim Pois(4)$$

a. Find the PMF of X;

Poisson Distribution: Examples

Example 24.2: Assume

$$X \sim Pois(4)$$

a. Find the PMF of X;

b. Find $\mathbb{P}(X = 2)$;

Poisson Distribution: Examples

Example 24.2: Assume

$$X \sim Pois(4)$$

a. Find the PMF of X;
b. Find $\mathbb{P}(X = 2)$;
c. Find $\mathbb{P}(3 \leq X < 5.5)$;

Poisson Distribution: Examples

Example 24.2: Assume

$$X \sim Pois(4)$$

a. Find the PMF of X;
b. Find $\mathbb{P}(X = 2)$;
c. Find $\mathbb{P}(3 \leq X < 5.5)$;
d. Find $F(3.1)$, where F is the CDF of X;

Poisson Distribution: Examples

Example 24.2: Assume

$$X \sim Pois(4)$$

a. Find the PMF of X;
b. Find $\mathbb{P}(X = 2)$;
c. Find $\mathbb{P}(3 \leq X < 5.5)$;
d. Find $F(3.1)$, where F is the CDF of X;
e. Sketch the graph of the CDF F.

Poisson Distribution: Examples

Example 24.3: Assume that the average number of traffic jams during a day at the Arshakunyats ave. is approximately 2.3. Assuming that for the number of traffic jams the Poisson model is appropriate, let X be the number of traffic jams today.

Poisson Distribution: Examples

Example 24.3: Assume that the average number of traffic jams during a day at the Arshakunyats ave. is approximately 2.3. Assuming that for the number of traffic jams the Poisson model is appropriate, let X be the number of traffic jams today.

a. So we assume $X \sim Pois(\lambda)$. Find λ;

Poisson Distribution: Examples

Example 24.3: Assume that the average number of traffic jams during a day at the Arshakunyats ave. is approximately 2.3. Assuming that for the number of traffic jams the Poisson model is appropriate, let X be the number of traffic jams today.

a. So we assume $X \sim Pois(\lambda)$. Find λ;

b. Write down the PMF of X;

Poisson Distribution: Examples

Example 24.3: Assume that the average number of traffic jams during a day at the Arshakunyats ave. is approximately 2.3. Assuming that for the number of traffic jams the Poisson model is appropriate, let X be the number of traffic jams today.

a. So we assume $X \sim Pois(\lambda)$. Find λ;
b. Write down the PMF of X;
c. Calculate the probability that no traffic jams will be today at the Arshakunyats ave.;

Poisson Distribution: Examples

Example 24.3: Assume that the average number of traffic jams during a day at the Arshakunyats ave. is approximately 2.3. Assuming that for the number of traffic jams the Poisson model is appropriate, let X be the number of traffic jams today.

a. So we assume $X \sim \mathit{Pois}(\lambda)$. Find λ;
b. Write down the PMF of X;
c. Calculate the probability that no traffic jams will be today at the Arshakunyats ave.;
d. Calculate the probability that exactly 2 traffic jams will be today at the Arshakunyats ave.;

Poisson Distribution: Examples

Example 24.3: Assume that the average number of traffic jams during a day at the Arshakunyats ave. is approximately 2.3. Assuming that for the number of traffic jams the Poisson model is appropriate, let X be the number of traffic jams today.

a. So we assume $X \sim Pois(\lambda)$. Find λ;
b. Write down the PMF of X;
c. Calculate the probability that no traffic jams will be today at the Arshakunyats ave.;
d. Calculate the probability that exactly 2 traffic jams will be today at the Arshakunyats ave.;
e. Calculate the probability that the number of traffic jams there will be more than 5.

Relation of Binomial and Poisson Distributions

It is important that the Poisson Distribution is the limiting case for Binomial Distribution, Binomial Probabilities.

Relation of Binomial and Poisson Distributions

It is important that the Poisson Distribution is the limiting case for Binomial Distribution, Binomial Probabilities. Recall that when doing n independent trials with probability p of success in each trial, then the Probability of having exactly k successes is (X is the number of successes):

$$\mathbb{P}(X = k)$$

Relation of Binomial and Poisson Distributions

It is important that the Poisson Distribution is the limiting case for Binomial Distribution, Binomial Probabilities. Recall that when doing n independent trials with probability p of success in each trial, then the Probability of having exactly k successes is (X is the number of successes):

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

Relation of Binomial and Poisson Distributions

It is important that the Poisson Distribution is the limiting case for Binomial Distribution, Binomial Probabilities. Recall that when doing n independent trials with probability p of success in each trial, then the Probability of having exactly k successes is (X is the number of successes):

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

Now, assume n is large enough, but also p is small enough, and $\lambda = n \cdot p$.

Relation of Binomial and Poisson Distributions

It is important that the Poisson Distribution is the limiting case for Binomial Distribution, Binomial Probabilities. Recall that when doing n independent trials with probability p of success in each trial, then the Probability of having exactly k successes is (X is the number of successes):

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

Now, assume n is large enough, but also p is small enough, and $\lambda = n \cdot p$. Then,

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Relation of Binomial and Poisson Distributions

It is important that the Poisson Distribution is the limiting case for Binomial Distribution, Binomial Probabilities. Recall that when doing n independent trials with probability p of success in each trial, then the Probability of having exactly k successes is (X is the number of successes):

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

Now, assume n is large enough, but also p is small enough, and $\lambda = n \cdot p$. Then,

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \approx e^{-\lambda} \cdot \frac{\lambda^k}{k!}.$$

Relation of Binomial and Poisson Distributions

It is important that the Poisson Distribution is the limiting case for Binomial Distribution, Binomial Probabilities. Recall that when doing n independent trials with probability p of success in each trial, then the Probability of having exactly k successes is (X is the number of successes):

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

Now, assume n is large enough, but also p is small enough, and $\lambda = n \cdot p$. Then,

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \approx e^{-\lambda} \cdot \frac{\lambda^k}{k!}.$$

Idea: OTB

Relation of Binomial and Poisson Distributions

So sometimes, when we solve a problem with a Binomial Distribution, with large number of trials and small probability of Success, we can approximate the probability having k successes by using Poisson Distribution.

[1] By Prokhorov, Le Cam, Barbour et al.

Relation of Binomial and Poisson Distributions

So sometimes, when we solve a problem with a Binomial Distribution, with large number of trials and small probability of Success, we can approximate the probability having k successes by using Poisson Distribution.
And you can use the Poisson approximation when dealing with a r.v. X, with $X \sim Binomial(n, p)$, where

[1] By Prokhorov, Le Cam, Barbour et al.

Relation of Binomial and Poisson Distributions

So sometimes, when we solve a problem with a Binomial Distribution, with large number of trials and small probability of Success, we can approximate the probability having k successes by using Poisson Distribution.

And you can use the Poisson approximation when dealing with a r.v. X, with $X \sim Binomial(n, p)$, where

- n is large;

[1] By Prokhorov, Le Cam, Barbour et al.

Relation of Binomial and Poisson Distributions

So sometimes, when we solve a problem with a Binomial Distribution, with large number of trials and small probability of Success, we can approximate the probability having k successes by using Poisson Distribution.

And you can use the Poisson approximation when dealing with a r.v. X, with $X \sim Binomial(n, p)$, where

- n is large;
- p is small.

[1] By Prokhorov, Le Cam, Barbour et al.

Relation of Binomial and Poisson Distributions

So sometimes, when we solve a problem with a Binomial Distribution, with large number of trials and small probability of Success, we can approximate the probability having k successes by using Poisson Distribution.

And you can use the Poisson approximation when dealing with a r.v. X, with $X \sim Binomial(n, p)$, where

- n is large;
- p is small.

In that case you can assume X is Poisson r.v. , $X \sim Pois(\lambda)$, with $\lambda = n \cdot p$.

[1] By Prokhorov, Le Cam, Barbour et al.

Relation of Binomial and Poisson Distributions

So sometimes, when we solve a problem with a Binomial Distribution, with large number of trials and small probability of Success, we can approximate the probability having k successes by using Poisson Distribution.

And you can use the Poisson approximation when dealing with a r.v. X, with $X \sim Binomial(n, p)$, where

- n is large;
- p is small.

In that case you can assume X is Poisson r.v. , $X \sim Pois(\lambda)$, with $\lambda = n \cdot p$.

Remark, 18+ : In fact, for the error of the approximation, the following is known[1]: if $X \sim Binom(n, p)$ and $Y \sim Pois(np)$, then for any set $A \subset \{0, 1, 2, 3, ...\}$,

$$|\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)| \leq \min\left(p, np^2\right).$$

[1] By Prokhorov, Le Cam, Barbour et al.

Example 24.4: Assume that the probability of having a broken egg in an egg-box is 0.5%. Tomorrow, one of the supermarkets will receive 1000 egg-boxes. What is the probability that exactly 6 out of that 1000 egg-boxes will contain broken eggs?

Solution:

a. Using the Binomial Model:

b. Using the Poisson Approximation:

Relation of Binomial and Poisson Distributions

But what I want to stress is that Poisson Distribution is not just an approximation to Binomial Distribution, is another approach of solving probability problems about calculation of number of rare events.

Relation of Binomial and Poisson Distributions

But what I want to stress is that Poisson Distribution is not just an approximation to Binomial Distribution, is another approach of solving probability problems about calculation of number of rare events.
So why is it more appropriate to use Poisson Distribution in many cases? Let me explain by giving examples.

Relation of Binomial and Poisson Distributions

Assume we consider a publisher, who publishes books, and sometimes, very rarely, you can find misprinted characters in that books.

Relation of Binomial and Poisson Distributions

Assume we consider a publisher, who publishes books, and sometimes, very rarely, you can find misprinted characters in that books. We want to model the Distribution of misprinted characters in a page for that publisher. Let X be the number of misprinted characters in a (random) page of a (random) book.

Relation of Binomial and Poisson Distributions

Assume we consider a publisher, who publishes books, and sometimes, very rarely, you can find misprinted characters in that books. We want to model the Distribution of misprinted characters in a page for that publisher. Let X be the number of misprinted characters in a (random) page of a (random) book.

We will give two models: the first one is Binomial, and the second one is Poisson, and make comparison.

Relation of Binomial and Poisson Distributions

The Binomial Model: For the Binomial Model, we need to have

Relation of Binomial and Poisson Distributions

The Binomial Model: For the Binomial Model, we need to have

- the probability of having a character misprinted:

Relation of Binomial and Poisson Distributions

The Binomial Model: For the Binomial Model, we need to have

- the probability of having a character misprinted: we'll assume $p = \dfrac{1}{50000}$ (Statistics!);

Relation of Binomial and Poisson Distributions

The Binomial Model: For the Binomial Model, we need to have

- the probability of having a character misprinted: we'll assume $p = \dfrac{1}{50000}$ (Statistics!);
- and the number of characters in a page:

Relation of Binomial and Poisson Distributions

The Binomial Model: For the Binomial Model, we need to have

- the probability of having a character misprinted: we'll assume $p = \dfrac{1}{50000}$ (Statistics!);
- and the number of characters in a page: we'll assume $n = 3000$.

Relation of Binomial and Poisson Distributions

The Binomial Model: For the Binomial Model, we need to have

- the probability of having a character misprinted: we'll assume $p = \dfrac{1}{50000}$ (Statistics!);
- and the number of characters in a page: we'll assume $n = 3000$.
- independence in making misprints in characters in different positions.

Relation of Binomial and Poisson Distributions

The Binomial Model: For the Binomial Model, we need to have

- the probability of having a character misprinted: we'll assume $p = \dfrac{1}{50000}$ (Statistics!);
- and the number of characters in a page: we'll assume $n = 3000$.
- independence in making misprints in characters in different positions. Say, the event of having a misprint in the 3rd character in a page is independent of having misprinted the 10th character etc.

Relation of Binomial and Poisson Distributions

The Binomial Model: For the Binomial Model, we need to have

- the probability of having a character misprinted: we'll assume $p = \dfrac{1}{50000}$ (Statistics!);
- and the number of characters in a page: we'll assume $n = 3000$.
- independence in making misprints in characters in different positions. Say, the event of having a misprint in the 3rd character in a page is independent of having misprinted the 10th character etc.

Now, having all this, we can model this situation (the number of misprints in a page), using the Binomial Distribution:

Relation of Binomial and Poisson Distributions

We are making an independent repeated trials model:

Relation of Binomial and Poisson Distributions

We are making an independent repeated trials model:

- Trial $=$ considering if a character is misprinted or not;

Relation of Binomial and Poisson Distributions

We are making an independent repeated trials model:

- Trial $=$ considering if a character is misprinted or not;
- Success $=$ having a character misprinted;

Relation of Binomial and Poisson Distributions

We are making an independent repeated trials model:

- Trial $=$ considering if a character is misprinted or not;
- Success $=$ having a character misprinted;
- Probability of Success $= p = \frac{1}{50000}$;

Relation of Binomial and Poisson Distributions

We are making an independent repeated trials model:

- Trial $=$ considering if a character is misprinted or not;
- Success $=$ having a character misprinted;
- Probability of Success $= p = \frac{1}{50000}$;
- No of Trials $=$ No. of characters we consider $= n = 3000$;

Relation of Binomial and Poisson Distributions

We are making an independent repeated trials model:

- Trial = considering if a character is misprinted or not;
- Success = having a character misprinted;
- Probability of Success $= p = \frac{1}{50000}$;
- No of Trials = No. of characters we consider $= n = 3000$;

Then, under the Binomial model, the probability of having exactly k misprinted characters will be:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} =$$

Relation of Binomial and Poisson Distributions

We are making an independent repeated trials model:

- Trial = considering if a character is misprinted or not;
- Success = having a character misprinted;
- Probability of Success $= p = \frac{1}{50000}$;
- No of Trials = No. of characters we consider $= n = 3000$;

Then, under the Binomial model, the probability of having exactly k misprinted characters will be:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} =$$

$$= \binom{3000}{k} \cdot \frac{1}{50000^k} \cdot \left(1 - \frac{1}{50000}\right)^{3000-k}$$

Relation of Binomial and Poisson Distributions

For example, some values (rounded) are:

Values of X	0	1	2	3	...	3000
$\mathbb{P}(X = x)$	0.942	0.056	0.002	$3.4 \cdot 10^{-5}$	...	0

Relation of Binomial and Poisson Distributions

For example, some values (rounded) are:

Values of X	0	1	2	3	...	3000
$\mathbb{P}(X = x)$	0.942	0.056	0.002	$3.4 \cdot 10^{-5}$	...	0

Now, the point is that this will work for a page with $n = 3000$ characters. If we will have, say a page with 3200 characters, we need to recalculate everything.

Relation of Binomial and Poisson Distributions

The Poisson Model: Let's talk now about the Poisson Model for the same problem: in this model we consider **per page** misprints, the average number of misprints in a page.

Relation of Binomial and Poisson Distributions

The Poisson Model: Let's talk now about the Poisson Model for the same problem: in this model we consider **per page** misprints, the average number of misprints in a page. We assume that that average number of misprints in a page is $\lambda = 0.06$.

Relation of Binomial and Poisson Distributions

The Poisson Model: Let's talk now about the Poisson Model for the same problem: in this model we consider **per page** misprints, the average number of misprints in a page. We assume that that average number of misprints in a page is $\lambda = 0.06$. I have calculated that by multiplying *p*, the probability of a misprinted character by *n*, the number of characters in a page.

Relation of Binomial and Poisson Distributions

The Poisson Model: Let's talk now about the Poisson Model for the same problem: in this model we consider **per page** misprints, the average number of misprints in a page. We assume that that average number of misprints in a page is $\lambda = 0.06$. I have calculated that by multiplying p, the probability of a misprinted character by n, the number of characters in a page.

Well, in fact, people are not usually doing this way: usually people are using Statistics - in this case, we fix some books, we calculate the number of pages, we calculate the number of misprints in that pages, and divide the second one to the first one, to obtain the average number of misprints in a page.

Relation of Binomial and Poisson Distributions

So in the Poisson model we assume:

Relation of Binomial and Poisson Distributions

So in the Poisson model we assume:

- The average number of misprinted characters in a page is fixed, is $\lambda = 0.06$;

Relation of Binomial and Poisson Distributions

So in the Poisson model we assume:

- The average number of misprinted characters in a page is fixed, is $\lambda = 0.06$;
- misprints are independent.

Relation of Binomial and Poisson Distributions

So in the Poisson model we assume:

- The average number of misprinted characters in a page is fixed, is $\lambda = 0.06$;
- misprints are independent.

Now, under the Poisson Model, the Probability that we will have exactly k misprints will be:

$$\mathbb{P}(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!} = e^{-0.06} \cdot \frac{0.06^k}{k!}.$$

Relation of Binomial and Poisson Distributions

For example, some values (again rounded) are:

Values of X	0	1	2	3	...	3000	...
$\mathbb{P}(X = x)$	0.942	0.056	0.002	$3.4 \cdot 10^{-5}$	...	0	...

For example, some values (again rounded) are:

Values of X	0	1	2	3	$\dots$	3000	$\dots$
$\mathbb{P}(X = x)$	0.942	0.056	0.002	$3.4 \cdot 10^{-5}$	$\dots$	0	$\dots$

You can check that the (rounded) probabilities coincide for Binomial model and Poisson one: the exact values do not coincide, but are very close to each other:

Relation of Binomial and Poisson Distributions

Here I am giving these probabilities side by side:

Values of X	Binomial Prob	Poisson Prob
0	0.9417640	0.9417645
1	0.05650697	0.05650587
2	0.001694678	0.001695176
3	$3.387164e-05$	$3.390352e-05$
$\vdots$	$\vdots$	$\vdots$
3000	0	0
3001	$--$	0
3002	$--$	0
$\vdots$	$--$	0

Relation of Binomial and Poisson Distributions

Note: About the difference between distributions: you can see that Binomial Probabilities, for this example, are defined for $k = 0, ..., 3000$, but Poisson probabilities are defined for any $k = 0, 1, 2, ...$

Relation of Binomial and Poisson Distributions

Note: About the difference between distributions: you can see that Binomial Probabilities, for this example, are defined for $k = 0, ..., 3000$, but Poisson probabilities are defined for any $k = 0, 1, 2, ...$.

Note: Important difference in Binomial and Poisson Models: for the Poisson model, we forget/do not talk about the number of characters in a page!

Relation of Binomial and Poisson Distributions

Note: About the difference between distributions: you can see that Binomial Probabilities, for this example, are defined for $k = 0, ..., 3000$, but Poisson probabilities are defined for any $k = 0, 1, 2, ...$

Note: Important difference in Binomial and Poisson Models: for the Poisson model, we forget/do not talk about the number of characters in a page! The only important thing is the average number of misprints.

Relation of Binomial and Poisson Distributions

Another example:

Relation of Binomial and Poisson Distributions

Another example: assume we own a grocery store, and we want to be sure that we will have enough Parmigiano cheese triangles daily for our customers.

Relation of Binomial and Poisson Distributions

Another example: assume we own a grocery store, and we want to be sure that we will have enough Parmigiano cheese triangles daily for our customers. Possible approaches:

Relation of Binomial and Poisson Distributions

Another example: assume we own a grocery store, and we want to be sure that we will have enough Parmigiano cheese triangles daily for our customers. Possible approaches:

- we can order a large number of triangle-packaged Parmigianos, not to have a shortage.

Relation of Binomial and Poisson Distributions

Another example: assume we own a grocery store, and we want to be sure that we will have enough Parmigiano cheese triangles daily for our customers. Possible approaches:

- we can order a large number of triangle-packaged Parmigianos, not to have a shortage. Not preferable.

Relation of Binomial and Poisson Distributions

Another example: assume we own a grocery store, and we want to be sure that we will have enough Parmigiano cheese triangles daily for our customers. Possible approaches:

- we can order a large number of triangle-packaged Parmigianos, not to have a shortage. Not preferable.
- we can order a small number of cheese packages, "um chhasav - hairn el anitsats" $\ddot\smile$

Relation of Binomial and Poisson Distributions

Another example: assume we own a grocery store, and we want to be sure that we will have enough Parmigiano cheese triangles daily for our customers. Possible approaches:

- we can order a large number of triangle-packaged Parmigianos, not to have a shortage. Not preferable.
- we can order a small number of cheese packages, "um chhasav - hairn el anitsats" $\ddot\smile$ Not so preferable, either.

Relation of Binomial and Poisson Distributions

Another example: assume we own a grocery store, and we want to be sure that we will have enough Parmigiano cheese triangles daily for our customers. Possible approaches:

- we can order a large number of triangle-packaged Parmigianos, not to have a shortage. Not preferable.
- we can order a small number of cheese packages, "um chhasav - hairn el anitsats" $\ddot\smile$ Not so preferable, either.
- If we will do it in a little bit wise way, we can calculate the probability that in a day customers will buy more than, say, k packages.

Relation of Binomial and Poisson Distributions

Another example: assume we own a grocery store, and we want to be sure that we will have enough Parmigiano cheese triangles daily for our customers. Possible approaches:

- we can order a large number of triangle-packaged Parmigianos, not to have a shortage. Not preferable.
- we can order a small number of cheese packages, "um chhasav - hairn el anitsats" $\ddot\smile$ Not so preferable, either.
- If we will do it in a little bit wise way, we can calculate the probability that in a day customers will buy more than, say, k packages. Let X be the number of customers that will buy that cheese tomorrow.

Relation of Binomial and Poisson Distributions

Another example: assume we own a grocery store, and we want to be sure that we will have enough Parmigiano cheese triangles daily for our customers. Possible approaches:

- we can order a large number of triangle-packaged Parmigianos, not to have a shortage. Not preferable.
- we can order a small number of cheese packages, "um chhasav - hairn el anitsats" $\ddot\smile$ Not so preferable, either.
- If we will do it in a little bit wise way, we can calculate the probability that in a day customers will buy more than, say, k packages. Let X be the number of customers that will buy that cheese tomorrow. We will calculate $\mathbb{P}(X > k)$ for different values of k, and choose the minimum value of k such that $\mathbb{P}(X > k) < 0.05$ - so we will leave a small, 5%, chance (risk) that not all our clients will be satisfied, but that's OK for us.

Relation of Binomial and Poisson Distributions

So we need to calculate $\mathbb{P}(X > k)$.

Relation of Binomial and Poisson Distributions

So we need to calculate $\mathbb{P}(X > k)$. To that end, we need to have the distribution of X. We can do it in two ways:

So we need to calculate $\mathbb{P}(X > k)$. To that end, we need to have the distribution of X. We can do it in two ways:

Meth 1: We can use the number of possible customers tomorrow, n, and the probability of a customer to buy that cheese, p, and assume $X \sim Binom(n, p)$. Then calculate the above probabilities;

Relation of Binomial and Poisson Distributions

So we need to calculate $\mathbb{P}(X > k)$. To that end, we need to have the distribution of X. We can do it in two ways:

Meth 1: We can use the number of possible customers tomorrow, n, and the probability of a customer to buy that cheese, p, and assume $X \sim Binom(n, p)$. Then calculate the above probabilities;

Meth 2: We can calculate the average daily number of customers that buy that cheese, λ (Statistics!!!), and assume $X \sim Pois(\lambda)$. Then calculate the above probabilities.

Relation of Binomial and Poisson Distributions

So we need to calculate $\mathbb{P}(X > k)$. To that end, we need to have the distribution of X. We can do it in two ways:

Meth 1: We can use the number of possible customers tomorrow, n, and the probability of a customer to buy that cheese, p, and assume $X \sim Binom(n, p)$. Then calculate the above probabilities;

Meth 2: We can calculate the average daily number of customers that buy that cheese, λ (Statistics!!!), and assume $X \sim Pois(\lambda)$. Then calculate the above probabilities.

The second approach is usually what people are using.

Poisson Distribution: Examples

Example 24.5: Assume that the probability of having no misprinted character in a book page is 96% for some publisher.

Poisson Distribution: Examples

Example 24.5: Assume that the probability of having no misprinted character in a book page is 96% for some publisher.

a. Calculate the probability that we will find exactly 1 misprints in one page of that publisher's book;

Poisson Distribution: Examples

Example 24.5: Assume that the probability of having no misprinted character in a book page is 96% for some publisher.

a. Calculate the probability that we will find exactly 1 misprints in one page of that publisher's book;
b. Calculate the probability that we will find more than 3 misprints in one page of that publisher's book.

Poisson Distribution: Examples

Example 24.6:

a. Generate Poisson random numbers in **R**;
b. Plot in **R** and compare the PMFs of Poisson r.v. with different rates;
c. Plot the CDF of some Poisson r.v. in **R**.

Poisson Distribution Property

The Poisson Distribution has the following nice property: assume X is a r.v. counting the number of occurrences of some event during one unit of measure, say during 1 hour.

Poisson Distribution Property

The Poisson Distribution has the following nice property: assume X is a r.v. counting the number of occurrences of some event during one unit of measure, say during 1 hour. If we model X by a Poisson r.v., $X \sim Pois(\lambda)$, then λ is the hourly average number of that events.

Poisson Distribution Property

The Poisson Distribution has the following nice property: assume X is a r.v. counting the number of occurrences of some event during one unit of measure, say during 1 hour. If we model X by a Poisson r.v., $X \sim \mathit{Pois}(\lambda)$, then λ is the hourly average number of that events. Then we can calculate

Poisson Distribution Property

The Poisson Distribution has the following nice property: assume X is a r.v. counting the number of occurrences of some event during one unit of measure, say during 1 hour. If we model X by a Poisson r.v., $X \sim \mathit{Pois}(\lambda)$, then λ is the hourly average number of that events. Then we can calculate

- $\mathbb{P}(X = 0) =$ the probability of having 0 event in 1 hour;

Poisson Distribution Property

The Poisson Distribution has the following nice property: assume X is a r.v. counting the number of occurrences of some event during one unit of measure, say during 1 hour. If we model X by a Poisson r.v., $X \sim Pois(\lambda)$, then λ is the hourly average number of that events. Then we can calculate

- $\mathbb{P}(X = 0) =$ the probability of having 0 event in 1 hour;
- $\mathbb{P}(X = 1) =$ the probability of having 1 event in 1 hour;

Poisson Distribution Property

The Poisson Distribution has the following nice property: assume X is a r.v. counting the number of occurrences of some event during one unit of measure, say during 1 hour. If we model X by a Poisson r.v., $X \sim Pois(\lambda)$, then λ is the hourly average number of that events. Then we can calculate

- $\mathbb{P}(X = 0) =$ the probability of having 0 event in 1 hour;
- $\mathbb{P}(X = 1) =$ the probability of having 1 event in 1 hour;
- $\mathbb{P}(X = 2) =$ the probability of having 2 events in 1 hour, ...

Poisson Distribution Property

The Poisson Distribution has the following nice property: assume X is a r.v. counting the number of occurrences of some event during one unit of measure, say during 1 hour. If we model X by a Poisson r.v., $X \sim Pois(\lambda)$, then λ is the hourly average number of that events. Then we can calculate

- $\mathbb{P}(X=0)$ = the probability of having 0 event in 1 hour;
- $\mathbb{P}(X=1)$ = the probability of having 1 event in 1 hour;
- $\mathbb{P}(X=2)$ = the probability of having 2 events in 1 hour,

  ...

Now, assume we want to model the number of events in 3 hours, Y.

Poisson Distribution Property

The Poisson Distribution has the following nice property: assume X is a r.v. counting the number of occurrences of some event during one unit of measure, say during 1 hour. If we model X by a Poisson r.v., $X \sim Pois(\lambda)$, then λ is the hourly average number of that events. Then we can calculate

- $\mathbb{P}(X = 0) =$ the probability of having 0 event in 1 hour;
- $\mathbb{P}(X = 1) =$ the probability of having 1 event in 1 hour;
- $\mathbb{P}(X = 2) =$ the probability of having 2 events in 1 hour, ...

Now, assume we want to model the number of events in 3 hours, Y. Then, since λ is the hourly average number of events, we will have that the average number of events in 3 hours is

Poisson Distribution Property

The Poisson Distribution has the following nice property: assume X is a r.v. counting the number of occurrences of some event during one unit of measure, say during 1 hour. If we model X by a Poisson r.v., $X \sim Pois(\lambda)$, then λ is the hourly average number of that events. Then we can calculate

- $\mathbb{P}(X = 0) =$ the probability of having 0 event in 1 hour;
- $\mathbb{P}(X = 1) =$ the probability of having 1 event in 1 hour;
- $\mathbb{P}(X = 2) =$ the probability of having 2 events in 1 hour, ...

Now, assume we want to model the number of events in 3 hours, Y. Then, since λ is the hourly average number of events, we will have that the average number of events in 3 hours is 3λ.

Poisson Distribution Property

The Poisson Distribution has the following nice property: assume X is a r.v. counting the number of occurrences of some event during one unit of measure, say during 1 hour. If we model X by a Poisson r.v., $X \sim Pois(\lambda)$, then λ is the hourly average number of that events. Then we can calculate

- $\mathbb{P}(X = 0) =$ the probability of having 0 event in 1 hour;
- $\mathbb{P}(X = 1) =$ the probability of having 1 event in 1 hour;
- $\mathbb{P}(X = 2) =$ the probability of having 2 events in 1 hour,

  ...

Now, assume we want to model the number of events in 3 hours, Y. Then, since λ is the hourly average number of events, we will have that the average number of events in 3 hours is 3λ. So we will have

$$Y \sim Pois(3 \cdot \lambda).$$

Poisson Distribution: Examples

Example 24.7: Let X be the annual number of Atlantic hurricanes. Assume we can model X by the Poisson r.v.,

$$X \sim \mathit{Pois}(5.44).$$

a. What is the annual average number of Atlantic hurricanes?
b. What is the probability of having more than 10 Atlantic hurricanes in a year?
c. Let Y be the number of Atlantic hurricanes in 5 years. Give a model for Y;
d. What is the probability of having more that 30 hurricanes in 5 years?

Poisson Distribution: Examples

Example 24.8: Assume X is the hourly number of my LinkedIn profile views. Say, my profile has 2 views in 10 hours in average.

a. Write an appropriate Probabilistic model for X;

b. What is the probability that I will have 1 view in the next hour?

c. What is the probability that I will have 3 views in the next 5 hours?

d. Are the following events equiprobable:
   - $A =$ I will have 1 views in the next 1 hour;
   - $B =$ I will have 3 views in the next 3 hours?

Poisson Distribution: Examples

Example 24.9: Assume that X is the number of pizza orders at Pandok Yerevan between 1PM and 3PM; let X be a Poisson r.v. with rate 2. Find the distribution of the number of pizza orders at Pandok Yerevan between 1PM and 3PM for days, when there was at least one order.

Examples of data fitting using the Poisson distribution

Poisson Distribution: Bortkiewicz data

Example 24.10 This is a classical example of fitting data by the Poisson distribution. The data set is given by Ladislaus von Bortkiewicz, in 1898[2].

[2] https://archive.org/download/dasgesetzderklei00bortrich/dasgesetzderklei00

Poisson Distribution: Bortkiewicz data

Example 24.10 This is a classical example of fitting data by the Poisson distribution. The data set is given by Ladislaus von Bortkiewicz, in 1898[2]. It shows the annual number of Prussian cavalryman in 10 corps killed by the kick of a horse, for 20 years. Here is the data:

No of Deaths	0	1	2	3	4
No of Corps-year	109	65	22	3	1

Fit the Poisson Distribution to this data.

[2] https://archive.org/download/dasgesetzderklei00bortrich/dasgesetzderklei00

Poisson Distribution: R's Discoveries data

Example 24.11: There is a native dataset in **R**, called *Discoveries*. It shows the yearly number of important scientific discoveries in each year from 1860 to 1959.

Fit the Poisson Distribution to this data.

Poisson Distribution: Flying Bomb data

Example 24.12: This is another classic example. During the WWII, Germany army used unmanned aircraft, called V1 Flying Bombs, to attack London. They hit the ground at more or less random points. The area of London was divided into 576 sectors of about 1/4 km each, and the number of hits for each sector was later calculated, for 537 Flying Bombs. The number of hits was

No of Hits	0	1	2	3	4	≥ 5
No of Sectors	229	211	93	35	7	1

Fit the Poisson Distribution to this data (take ≥ 5 as 5).

Poisson Distribution: Raisins Problem

Example 24.13: Assume we have a 2D cookie in the form of a square, and raisins are randomly (and uniformly!) spread out around the cookie dough. Do the simulation in **R**, divide the cookie area into small squares by a square grid, calculate the number of raisins in each square, and fit the Poisson Distribution to the data for the number of raisins per small square.

Of course, this is the simulation of the previous, Flying Bombs, problem.

Poisson Distribution: Raisins Problem

Example 24.13: Assume we have a 2D cookie in the form of a square, and raisins are randomly (and uniformly!) spread out around the cookie dough. Do the simulation in **R**, divide the cookie area into small squares by a square grid, calculate the number of raisins in each square, and fit the Poisson Distribution to the data for the number of raisins per small square.
Of course, this is the simulation of the previous, Flying Bombs, problem. Sometimes it is formulated as calculating the number of raindrops (dust particles, atomic particles) over a small grid.

Poisson Distribution: Raisins Problem

Example 24.13: Assume we have a 2D cookie in the form of a square, and raisins are randomly (and uniformly!) spread out around the cookie dough. Do the simulation in **R**, divide the cookie area into small squares by a square grid, calculate the number of raisins in each square, and fit the Poisson Distribution to the data for the number of raisins per small square.
Of course, this is the simulation of the previous, Flying Bombs, problem. Sometimes it is formulated as calculating the number of raindrops (dust particles, atomic particles) over a small grid.

Exercise: State and do 1D and 3D case problems and simulations.

Poisson Distribution: Hurricane data

Example 24.14: At `http://www.stormfax.com/huryear.htm` you can find the Atlantic Hurricane Numbers By Year for 1851-2017.

Fit the Poisson Distribution to the data for the number of Atlantic Hurricanes.

Poisson Distribution: Hurricane data

Example 24.15: You can find data for the distribution of word lengths in various languages at `http://www.ravi.io/language-word-lengths`.

Fit the Poisson Distribution to the data for the number of characters for German words.

Poisson Distribution: Goals data

Example 24.16: You can find the UK Premiere League matches results for 2019/2020 at `https://www.football-data.co.uk/englandm.php` .

Fit the Poisson Distribution to the data for the number of goals.

Negative Binomial Distribution