CS 107, Probability, Spring 2019
Lecture 24

Michael Poghosyan

AUA

20 March 2019

Content

- Examples of Important Discrete R.V.s

MidSem Evaluation Results

Overview: Questions: 7, Feedback: 11

MidSem Evaluation Results

Overview: Questions: 7, Feedback: 11

Multiple Choice Questions $\ddot\smile$

MidSem Evaluation Results

Overview: Questions: 7, Feedback: 11

Multiple Choice Questions $\ddot\smile$

- Q: Are you using our textbooks? No: 1, Yes: 9

MidSem Evaluation Results

Overview: Questions: 7, Feedback: 11

Multiple Choice Questions $\ddot\smile$

- Q: Are you using our textbooks? No: 1, Yes: 9
- Q: Do we need a LZ section for each lecture? No: 1, Yes: 10

MidSem Evaluation Results

Overview: Questions: 7, Feedback: 11

Multiple Choice Questions $\ddot\smile$

- Q: Are you using our textbooks? No: 1, Yes: 9
- Q: Do we need a LZ section for each lecture? No: 1, Yes: 10
- Q: Are Quizzes adequate to the topics covered? No: 1, Yes: 10

MidSem Evaluation Results

Overview: Questions: 7, Feedback: 11

Multiple Choice Questions $\ddot\smile$

- Q: Are you using our textbooks? No: 1, Yes: 9
- Q: Do we need a LZ section for each lecture? No: 1, Yes: 10
- Q: Are Quizzes adequate to the topics covered? No: 1, Yes: 10
- Q: Are Homework adequate to the topics covered? No: 1, Yes: 10

MidSem Evaluation Results

Overview: Questions: 7, Feedback: 11

Multiple Choice Questions $\ddot\smile$

- Q: Are you using our textbooks? No: 1, Yes: 9
- Q: Do we need a LZ section for each lecture? No: 1, Yes: 10
- Q: Are Quizzes adequate to the topics covered? No: 1, Yes: 10
- Q: Are Homework adequate to the topics covered? No: 1, Yes: 10
- Q: Course load is: Veery light:0, Light: 0, OK: 5, A little bit heavy: 4, Heavy: 1

MidSem Evaluation Results

Open-Ended Question: What is NOT working well in this class?

MidSem Evaluation Results

Open-Ended Question: What is NOT working well in this class?

- For me when you write on the white board you erase it immediately and there is no time to write that down :(

MidSem Evaluation Results

Open-Ended Question: What is NOT working well in this class?

- For me when you write on the white board you erase it immediately and there is no time to write that down :(
- please don't clean the blackboard so fast, we don't manage to write everything down :(

MidSem Evaluation Results

Open-Ended Question: What is NOT working well in this class?

- For me when you write on the white board you erase it immediately and there is no time to write that down :(
- please don't clean the blackboard so fast, we don't manage to write everything down :(
- Having review of material after quiz and hw was not really helping. :) But overall it is great. Will miss your funny and interesting lecture notes

MidSem Evaluation Results

Open-Ended Question: What is NOT working well in this class?

MidSem Evaluation Results

Open-Ended Question: What is NOT working well in this class?

- The only problem for me is that we don't solve some types of problems in class. Meaning that for example we are solving problem with picking one ball from a box, but in homework we got a problem of picking 3 balls. This is very simple example, but I mean we have not solve in class the problems for 2 and more balls picking before we got the homework. Sometimes we solve such a problems only after the homework.

MidSem Evaluation Results

Open-Ended Question: Your suggestions to improve this course:

MidSem Evaluation Results

Open-Ended Question: Your suggestions to improve this course:

- if possible have notes because they are really helpful

MidSem Evaluation Results

Open-Ended Question: Your suggestions to improve this course:

- if possible have notes because they are really helpful
- I hope that we will continue having notes as even if we missed something during class we could learn it from notes. I think it would have been better if we had notes instead of slides.

MidSem Evaluation Results

Open-Ended Question: Your suggestions to improve this course:

- if possible have notes because they are really helpful
- I hope that we will continue having notes as even if we missed something during class we could learn it from notes. I think it would have been better if we had notes instead of slides.
- You are one of the best professors in the AUA right now. Hope you can share your ways of teaching with other professors. Although I am a student, that would be great if I will be able to attend it to learn skills I might need one day.

MidSem Evaluation Results

Open-Ended Question: Your suggestions to improve this course:

- if possible have notes because they are really helpful
- I hope that we will continue having notes as even if we missed something during class we could learn it from notes. I think it would have been better if we had notes instead of slides.
- You are one of the best professors in the AUA right now. Hope you can share your ways of teaching with other professors. Although I am a student, that would be great if I will be able to attend it to learn skills I might need one day. 

MidSem Evaluation Results

Open-Ended Question: Your suggestions to improve this course:

MidSem Evaluation Results

Open-Ended Question: Your suggestions to improve this course:

- Increase the pressure on students to make them to use textbooks as well. We are concerned with the lack of time for during some extra exercises from textbook, but they are necessary. So please, either pressure students to use textbook, or give at least 1 or 2 problems from textbook in the homework, so that we will be supposed to open that book. Thank you

MidSem Evaluation Results

Open-Ended Question: Your suggestions to improve this course:

- Increase the pressure on students to make them to use textbooks as well. We are concerned with the lack of time for during some extra exercises from textbook, but they are necessary. So please, either pressure students to use textbook, or give at least 1 or 2 problems from textbook in the homework, so that we will be supposed to open that book. Thank you

LZ

Geometric Distribution

Geometric Distribution

We will say that the r.v. X has a Geometric Distribution with the parameter p, $p \in (0, 1]$, and we will write $X \sim Geom(p)$, if its PMF is given by

$$\mathbb{P}(X = x) = p \cdot (1 - p)^{x-1}, \qquad x \in \mathbb{N}$$

that is,

Values of X	1	2	3	...	k	...
$\mathbb{P}(X = x)$	p	$p(1-p)$	$p(1-p)^2$	...	$p(1-p)^{k-1}$	...

Geometric Distribution

Geometric Distribution

We will say that the r.v. X has a Geometric Distribution with the parameter p, $p \in (0,1]$, and we will write $X \sim Geom(p)$, if its PMF is given by

$$\mathbb{P}(X = x) = p \cdot (1-p)^{x-1}, \qquad x \in \mathbb{N}$$

that is,

Values of X	1	2	3	$\dots$	k	$\dots$
$\mathbb{P}(X = x)$	p	$p(1-p)$	$p(1-p)^2$	$\dots$	$p(1-p)^{k-1}$	$\dots$

- **Exercise:** Prove that the probabilities sum up to 1.

Geometric Distribution

Geometric Distribution

We will say that the r.v. X has a Geometric Distribution with the parameter p, $p \in (0, 1]$, and we will write $X \sim Geom(p)$, if its PMF is given by

$$\mathbb{P}(X = x) = p \cdot (1 - p)^{x-1}, \qquad x \in \mathbb{N}$$

that is,

Values of X	1	2	3	$\dots$	k	$\dots$
$\mathbb{P}(X = x)$	p	$p(1-p)$	$p(1-p)^2$	$\dots$	$p(1-p)^{k-1}$	$\dots$

- **Exercise:** Prove that the probabilities sum up to 1.
- **Exercise:** Calculate the CDF of $X \sim Geom(p)$.

Geometric Distribution: Interpretation

Geometric Distribution gives the model for repeating trials until the first Success.

Geometric Distribution: Interpretation

Geometric Distribution gives the model for repeating trials until the first Success.

If $X \sim Geom(p)$, then

Geometric Distribution gives the model for repeating trials until the first Success.

If $X \sim Geom(p)$, then

- we are doing independent repetitions of some Simple Experiment,

Geometric Distribution: Interpretation

Geometric Distribution gives the model for repeating trials until the first Success.

If $X \sim Geom(p)$, then

- we are doing independent repetitions of some Simple Experiment,
- in each Simple Experiment some Event can happen (Success) with the Probability p;

Geometric Distribution: Interpretation

Geometric Distribution gives the model for repeating trials until the first Success.

If $X \sim Geom(p)$, then

- we are doing independent repetitions of some Simple Experiment,
- in each Simple Experiment some Event can happen (Success) with the Probability p;
- X shows the number of trial we will have until the first Success will be shown

Geometric Distribution: Note

Recall that the definition of the Geometric Distribution was:

Values of X	1	2	3	...	k	...
$\mathbb{P}(X = x)$	p	$p(1-p)$	$p(1-p)^2$	...	$p(1-p)^{k-1}$	...

Geometric Distribution: Note

Recall that the definition of the Geometric Distribution was:

Values of X	1	2	3	...	k	...
$\mathbb{P}(X = x)$	p	$p(1-p)$	$p(1-p)^2$	...	$p(1-p)^{k-1}$	...

X is counting **the number of trial we will have until the first Success**.

Geometric Distribution: Note

Recall that the definition of the Geometric Distribution was:

Values of X	1	2	3	...	k	...
$\mathbb{P}(X = x)$	p	$p(1-p)$	$p(1-p)^2$	...	$p(1-p)^{k-1}$	...

X is counting **the number of trial we will have until the first Success**.

In some books, people define the Geometric Distribution in the other way, starting from the value 0:

Values of X	0	1	2	...	k	...
$\mathbb{P}(X = x)$	p	$p(1-p)$	$p(1-p)^2$	...	$p(1-p)^k$	...

Geometric Distribution: Note

Recall that the definition of the Geometric Distribution was:

Values of X	1	2	3	...	k	...
$\mathbb{P}(X = x)$	p	$p(1-p)$	$p(1-p)^2$	...	$p(1-p)^{k-1}$	...

X is counting **the number of trial we will have until the first Success**.

In some books, people define the Geometric Distribution in the other way, starting from the value 0:

Values of X	0	1	2	...	k	...
$\mathbb{P}(X = x)$	p	$p(1-p)$	$p(1-p)^2$	...	$p(1-p)^k$	...

In this case X shows **the number of Failures before the first Success**.

Geometric Distribution: Examples

- Assume we are rolling a die several times until first time we will get 5 shown. Then p is the probability to get 5 in one roll, and we know that for a fair die $p = \frac{1}{6}$. Let X be the number of roll that will show 5 first time. Then $X \sim Geom(\frac{1}{6})$.

Geometric Distribution: Examples

- Assume we are rolling a die several times until first time we will get 5 shown. Then p is the probability to get 5 in one roll, and we know that for a fair die $p = \frac{1}{6}$. Let X be the number of roll that will show 5 first time. Then $X \sim Geom(\frac{1}{6})$.
- Assume a couple wants to calculate the number of children they will have until the first boy. Assume also we know the probability of having a boy, $p = 0.49$, and suppose that having a boy or girl is independent of the previous children. Let X be the number of children they will have until the first boy (say, $X = 1$ means that the first child was a boy, $X = 3$ means the first 2 children were girls and 3rd one was a boy etc.). Then $X \sim Geom(0.49)$.

Geometric Distribution: Examples

- Say, we are asking randomly AUA students until we will find a GPA 4.0 student. Then the number of students we will ask until meeting a GPA 4.0 student is $X \sim Geom(p)$, where p is the probability of AUA student to be a 4.0 GPA student (the ratio of 4.0 GPA students to all AUA students).

Geometric Distribution: Examples

- Say, we are asking randomly AUA students until we will find a GPA 4.0 student. Then the number of students we will ask until meeting a GPA 4.0 student is $X \sim Geom(p)$, where p is the probability of AUA student to be a 4.0 GPA student (the ratio of 4.0 GPA students to all AUA students). BTW, this can give you a method to estimate the proportion (i.e., the probability p) of AUA 4.0 GPA students. Just go out and randomly ask students until getting the first 4.0 GPA-shnik. Then use some statistics to estimate p $\ddot\smile$

Geometric Distribution: Examples

- Say, we are asking randomly AUA students until we will find a GPA 4.0 student. Then the number of students we will ask until meeting a GPA 4.0 student is $X \sim Geom(p)$, where p is the probability of AUA student to be a 4.0 GPA student (the ratio of 4.0 GPA students to all AUA students). BTW, this can give you a method to estimate the proportion (i.e., the probability p) of AUA 4.0 GPA students. Just go out and randomly ask students until getting the first 4.0 GPA-shnik. Then use some statistics to estimate p $\ddot\smile$
- Your turn!

Geometric Distribution: R Examples

R Code

```
#Geomteric Distribution
p = 0.7
smpl <- rgeom(20, prob = p)
smpl

#Plotting the PMF
n = 10
x <- 0:n
PMF <- dgeom(x, prob = p)
plot(x,PMF, pch = 19)
#Now, the CDF
t <- seq(-2,5, 0.1)
CDF <- pgeom(t, p)
plot(t, CDF, type = "s")
```

Poisson Distribution

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

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

that is,

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

Poisson Distribution

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

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

that is,

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

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

Poisson Distribution: Interpretation

Poisson Distribution is modeling the number of times an event occurs in a fixed interval of time or space, and λ is the rate of that event. For example, we can model the following events by using the Poisson distribution:

Poisson Distribution: Interpretation

Poisson Distribution is modeling the number of times an event occurs in a fixed interval of time or space, and λ is the rate of that event. For example, we can model the following events by using the Poisson distribution:

- The number of emails received in a day

Poisson Distribution: Interpretation

Poisson Distribution is modeling the number of times an event occurs in a fixed interval of time or space, and λ is the rate of that event. 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: Interpretation

Poisson Distribution is modeling the number of times an event occurs in a fixed interval of time or space, and λ is the rate of that event. 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 between 9AM and 9PM

Poisson Distribution: Interpretation

Poisson Distribution is modeling the number of times an event occurs in a fixed interval of time or space, and λ is the rate of that event. 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 between 9AM and 9PM
- The number of meteors greater than 1 meter diameter that strike earth in a year (Wiki)

Poisson Distribution: Interpretation

Poisson Distribution is modeling the number of times an event occurs in a fixed interval of time or space, and λ is the rate of that event. 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 between 9AM and 9PM
- The number of meteors greater than 1 meter diameter that strike earth in a year (Wiki)
- The number of occurrences of the DNA sequence "ACGT" in a gene (Wiki)

Poisson Distribution: Interpretation

Poisson Distribution is modeling the number of times an event occurs in a fixed interval of time or space, and λ is the rate of that event. 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 between 9AM and 9PM
- The number of meteors greater than 1 meter diameter that strike earth in a year (Wiki)
- The number of occurrences of the DNA sequence "ACGT" in a gene (Wiki)
- The number of patients arriving in an emergency room between 11 and 12 pm (Wiki)

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 during which we are considering our events (say, every minute, every day, every month, every day from 10:30AM till 11:20AM 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 during which we are considering our events (say, every minute, every day, every month, every day from 10:30AM till 11:20AM 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 during which we are considering our events (say, every minute, every day, every month, every day from 10:30AM till 11:20AM 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 during which we are considering our events (say, every minute, every day, every month, every day from 10:30AM till 11:20AM 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 during which we are considering our events (say, every minute, every day, every month, every day from 10:30AM till 11:20AM 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

And in the above case, λ is the average number of events in the interval, or the rate of the events.

Poisson Distribution: Examples

And in the above case, λ is the average number of events in the interval, or the rate of the events.
Some examples:

- (Wiki) The number of students who arrive at the student union per minute will likely not follow a Poisson distribution, because the rate is not constant (low rate during class time, high rate between class times) and the arrivals of individual students are not independent (students tend to come in groups).

Poisson Distribution: Examples

And in the above case, λ is the average number of events in the interval, or the rate of the events.

Some examples:

- (Wiki) The number of students who arrive at the student union per minute will likely not follow a Poisson distribution, because the rate is not constant (low rate during class time, high rate between class times) and the arrivals of individual students are not independent (students tend to come in groups).
- The number of students absent from our Probability course.

Poisson Distribution: Examples

And in the above case, λ is the average number of events in the interval, or the rate of the events.
Some examples:

- (Wiki) The number of students who arrive at the student union per minute will likely not follow a Poisson distribution, because the rate is not constant (low rate during class time, high rate between class times) and the arrivals of individual students are not independent (students tend to come in groups).
- The number of students absent from our Probability course. This number, X, is a r.v.

Poisson Distribution: Examples

And in the above case, λ is the average number of events in the interval, or the rate of the events.
Some examples:

- (Wiki) The number of students who arrive at the student union per minute will likely not follow a Poisson distribution, because the rate is not constant (low rate during class time, high rate between class times) and the arrivals of individual students are not independent (students tend to come in groups).
- The number of students absent from our Probability course. This number, X, is a r.v. But it is not a Poisson r.v., since the rate is not the same for all days - on Fridays, we have less absents (because of HW and Quiz $\ddot\smile$).

Poisson Distribution: Examples

And in the above case, λ is the average number of events in the interval, or the rate of the events.
Some examples:

- (Wiki) The number of students who arrive at the student union per minute will likely not follow a Poisson distribution, because the rate is not constant (low rate during class time, high rate between class times) and the arrivals of individual students are not independent (students tend to come in groups).
- The number of students absent from our Probability course. This number, X, is a r.v. But it is not a Poisson r.v., since the rate is not the same for all days - on Fridays, we have less absents (because of HW and Quiz $\ddot\smile$). Can you suggest a method to make it a Poisson r.v.?

Poisson Distribution: Examples

Poisson Distribution: Examples

- Assume that the average number of traffic jams during a day at the Arshakunyats street is approximately 2.3. Here the Poisson model is appropriate, and since the average event rate is 2.3 jams per day, $\lambda = 2.3$.

Poisson Distribution: Examples

- Assume that the average number of traffic jams during a day at the Arshakunyats street is approximately 2.3. Here the Poisson model is appropriate, and since the average event rate is 2.3 jams per day, $\lambda = 2.3$.
  For this example, we can calculate the probabilities like

$$\mathbb{P}(\text{exactly 0 jams will happen today}) = \frac{2.3^0}{0!}e^{-2.3} = e^{-2.3}$$

$$\mathbb{P}(\text{exactly 1 jam will happen today}) = \frac{2.3^1}{1!}e^{-2.3} = 2.3 \cdot e^{-2.3}$$

$$\mathbb{P}(\text{at most 5 jams will happen today}) = \sum_{k=0}^{5} \frac{2.3^k}{k!}e^{-2.3}$$

etc.

Poisson Distribution: Examples

Poisson Distribution: Examples

- Classic Poisson example: the data set of von Bortkiewicz (1898), for the chance of a Prussian cavalryman being killed by the kick of a horse. See, for example, `http://www.umass.edu/wsp/resources/poisson/index.html`

Poisson Distribution: Examples

- Classic Poisson example: the data set of von Bortkiewicz (1898), for the chance of a Prussian cavalryman being killed by the kick of a horse. See, for example, `http://www.umass.edu/wsp/resources/poisson/index.html`
- DeepMind Lecture on Reinforcement Learning!

Poisson Distribution: Examples

- Classic Poisson example: the data set of von Bortkiewicz (1898), for the chance of a Prussian cavalryman being killed by the kick of a horse. See, for example, `http://www.umass.edu/wsp/resources/poisson/index.html`
- DeepMind Lecture on Reinforcement Learning!
- Some other examples: the number of visitors to a web site per minute; the number of typographical errors found in a book; the number of network failures per day; the number of my OH visitors (maybe no? this changes close to midterms ?); the number of hungry professors visiting AUA cafeteria between 12:30PM and 14:30PM on MWF;

Poisson Distribution: Examples

- Classic Poisson example: the data set of von Bortkiewicz (1898), for the chance of a Prussian cavalryman being killed by the kick of a horse. See, for example, `http://www.umass.edu/wsp/resources/poisson/index.html`
- DeepMind Lecture on Reinforcement Learning!
- Some other examples: the number of visitors to a web site per minute; the number of typographical errors found in a book; the number of network failures per day; the number of my OH visitors (maybe no? this changes close to midterms ?); the number of hungry professors visiting AUA cafeteria between 12:30PM and 14:30PM on MWF;
- Your turn!

Poisson Distribution: R Examples

R Code

```
#Poisson Distribution
lambda <- 1.5
n <- 10
x <- 0:n
x.prob <- dpois(x,lambda) #PMF

#To have 2 plots side by side
par(mfrow=c(1,2))
#Plotting the PMF
plot(x, x.prob, type = "h", lwd = 5, main = "Poisson Di
#Now plotting the CDF
t <- seq(from = -1, to = n+1, by = 0.01)
y <- ppois(t,lambda)
plot(t,y, type = "s", lwd = 3, main = "Poisson Distribu
```

Approximation of Binomial Distribution through Poisson

Now, let us talk about one application of the Poisson Distribution for calculation of Binomial Probabilities.

Approximation of Binomial Distribution through Poisson

Now, let us talk about one application of the Poisson Distribution for calculation of 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):

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

Approximation of Binomial Distribution through Poisson

Now, let us talk about one application of the Poisson Distribution for calculation of 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):

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

Approximation of Binomial Distribution through Poisson

Now, let us talk about one application of the Poisson Distribution for calculation of 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 $\lambda = n \cdot p$.

Approximation of Binomial Distribution through Poisson

Now, let us talk about one application of the Poisson Distribution for calculation of 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,

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

Approximation of Binomial Distribution through Poisson

Now, let us talk about one application of the Poisson Distribution for calculation of 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,

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

Poisson Distribution: R Examples

R Code

```
# Approximation of Binomial by Poisson
n <- 100
p <- 0.004
lambda <- n*p

x <- 0:20
y <- dbinom(x, size = n, prob = p)
plot(x,y, ylim = c(0,0.4))
z <- dpois(x, lambda = lambda)
par(new=T)
plot(x,z, pch = 18, col = "red", ylim = c(0,0.4))

#Plotting the difference
plot(x, abs(y-z))
```

Poisson Distribution: R Examples

R Code

```
##### Real-life Example
help("discoveries")
disc <- discoveries
disc
plot(disc)
hist(disc, breaks = seq(0,13,1))

#Fitting the data by the Poisson
lambda = mean(disc)
table(disc)
n = 0:15
m = dpois(n, lambda)
plot(n,m, xlim = c(0,14), ylim = c(0,0.3), col="red",
par(new = T)
```

Poisson Distribution: R Examples

R Code

```
#Bortkiewicz's Data
#Deaths by horse-kick in Prussian Army cavalry corps
#(Bortkiewicz 1898). N is the number of units
#corresponding to each number of deaths.
hk<-data.frame(D = c(0,1,2,3,4,5), N = c(109,65,22,3,1,0))
plot(hk)
lambda <- sum(hk$D*hk$N)/sum(hk$N)

n = 0:5
m = dpois(n, lambda)
plot(n,m, xlim = c(0,5), ylim = c(0,0.6), col="red", type =
par(new = T)
plot(hk$D, hk$N/sum(hk$N) , xlim = c(0,5), ylim = c(0,0.6)
```