Distributions are following from this book:

* Probability and Statistics for Computer Scientists (3rd Edition) – Michael Baron

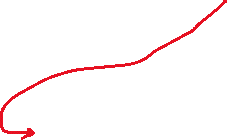
Bernoulli Distribution

Definition 3.10:

* A discrete random variable with 2 possible values, 0 and 1, is called a Bernoulli variable,
* its distribution is Bernoulli distribution,
* and any experiment with a *binary outcome* is called a Bernouilli trial.
* Variables are independent.
* We just try once.

Graphical user interface, text

Description automatically generated



We can use f(x) also. This is distribution function.

Since they are independent, we can say q=1-p

You will be given an experiment. You will decide what kind of event it is. Then you will define probability function. Then you will find expected value and variance out of that distribution. This is one question in exam.

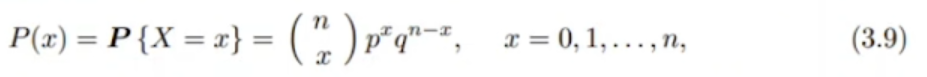
You need to know what it is. Computation is not much problem.

Binomial Distribution

Definition 3.11:

* A variable described as the # of successes in a sequence of independent Bernouilli trials has Binomial distribution.
* Its parameters are n, the # of trials, and p, the probability of success.
* We try many times.

Binomial probability mass function is:



We have bc we have certain number of success. Even if you don’t know this distribution, if you are asked that how many times you can get tail out of n trials, you would still use this combination.

A picture containing graphical user interface

Description automatically generated

x is # of success. Discrete random variable.

Geometric Distribution

Definition 3.12:

* The # of Bernouilli trials needed to get the first success has Geometric distribution.
* Once you get the success, you stop.

Graphical user interface

Description automatically generated

We have x trials. is 1 so we don’t need here.

Negative Binomial Distribution

Definition 3.13:

* In a sequence of independent Bernouilli trials, the # of trials needed to obtain k successes has Negative Binomial distribution.

In some sense, Negative Binomial distribution is opposite to Binomial distribution. Binomial variables count the # of successes in a fixed # of trials whereas Negative Binomial variables count the # of trials needed to see a fixed (k) # of successes. Other than this, there is nothing “negative” about this distribution.

# of trials is unknown. This is our variable x.

Negative Binomial probability mass function is

Text, letter

Description automatically generated

This formula accounts for the probability of k successes, the remaining (x-k) failures, and the # of outcomes-sequences with the k-th success coming on the x-th trial.

x-k 🡪 # of failure.

Negative Binomial distribution has 2 parameters, k (# of success we want) and p (probability of success).

With k = 1, it becomes Geometric.

Also, each Negative Binomial variable can be represented as a sum of independent Geometric variables,



with the same probability of success p.

Indeed, the # of trials until the k-th success consists of a Geometric number of trials X1 until the 1st success, an additional Geometric number of trials X2 until the 2nd success, etc.

For each success, we consider the event as one geometric distribution.

* You flip the coin until the first success and it is geometric.
* Then you start flipping the coin again and stop at the 2nd success.
* …
* You do that until the k-th success. So X is sum of k geometric distributions.

Because of 3.11, we have:

Chart, scatter chart

Description automatically generated

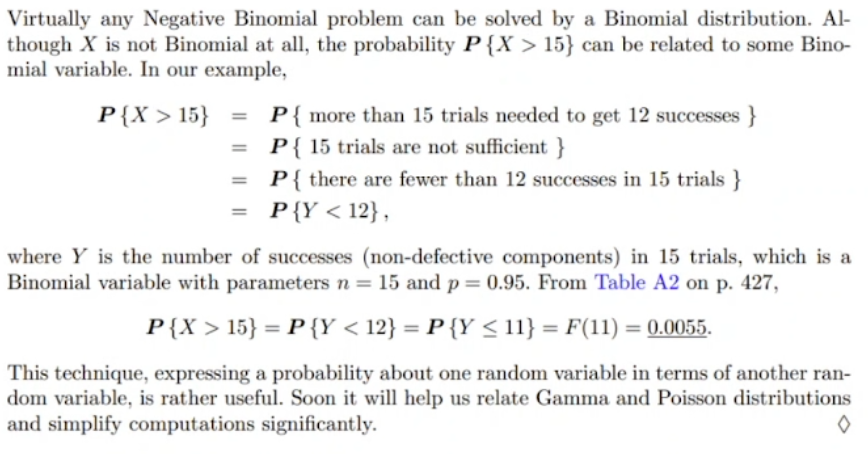
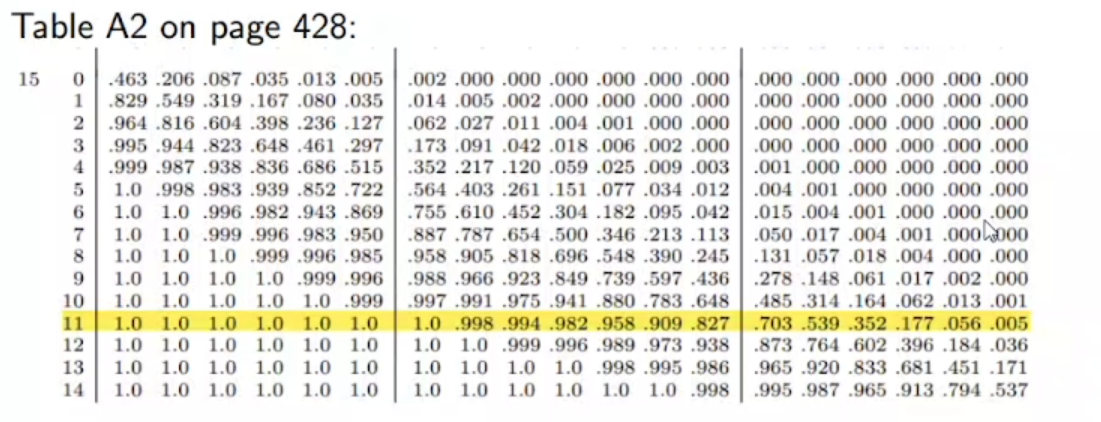
Distributions are not completely independent from each other. They are building on the same base.

Graphical user interface, application

Description automatically generated

Example (Sequential Testing):

* In a recent production, 5% of certain electronic components are defective.
* We need to find 12 non-defective components for our 12 new computers.
* Components are tested until 12 non-defective ones are found.
* What is the probability that more than 15 components will have to be tested?
* Graphical user interface, text

  Description automatically generated
* We need 12 successes so it is negative binomial.
* F(15) is actually f(15).
* 1 – F(15) 🡪 cumulative function until 15.
* 
* Y is number of success (non-defective components) in 15 trials.
* So in our book table, n is 15. It is # of trials. We know it.
* # of successes becomes the variable now. It should be less than 12.

Poisson Distribution

Definition 3.14:

* The # of rare events occurring within a fixed period of time has Poisson distribution.

This distribution bears the name of a famous French mathematician Simeon-Denis Poisson (1781-1840).

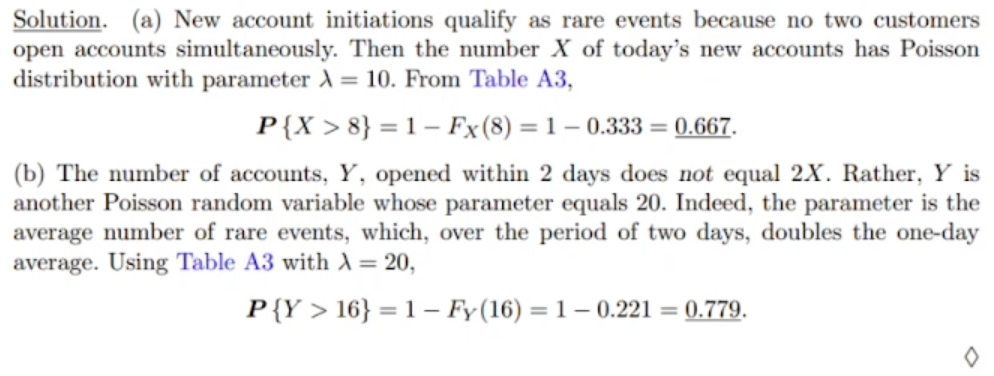
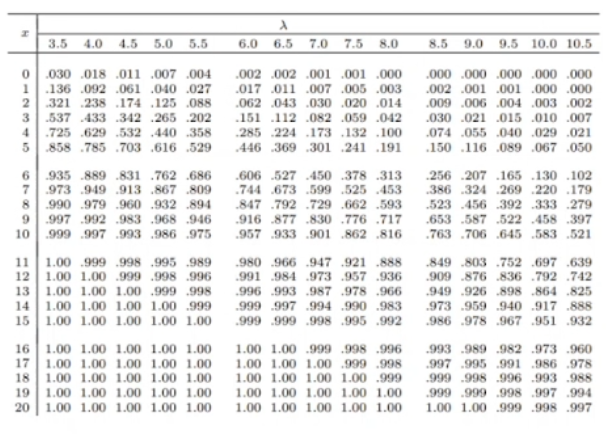
Text

Description automatically generated

Rare event 🡪 in a big amount of time, it happens less than certain amount of numbers

If we know frequency, we will expect that # of events in a certain amount of time. This is why we see in expected value.

Example (New Accounts):

* Customers of an internet service provider initiate new accounts at the average rate of 10 accounts per day.
  + (a) What is the probability that more than 8 new accounts will be initiated today?
  + (b) What is the probability that more than 16 accounts will be initiated within 2 days?
* 
* Rare event could be creating 2 events at the same time.
* Düzenli ve sürekli bir şekilde olmasını beklemediğimiz eventler rare eventlerdir.
* Average rate kelimeleri ile poisson kullanacağını anlarsın. Ayrıca fail or success durumu da yok.
* Rare ile kastedilen az oluyor değil, sadece belirli bir şeyde oluyor günde sadece 10 tane açılıyor.
* 
* Bir günde ortalama olarak 10 hesap açılacağını biliyorsun ki bu 8’den büyük. Yani (a) şıkkının 0.5’ten büyük olacağını soruyu okuduğunda tahmin edebilirsin.
* Frekansa ne kadar yakınsam, o kadar 1’e yakın bir olasılık beklerim.
* x <= olsun diye 1 – Fx(8) dedik ki tabloda kullanabilelim.
  + F ile gösterdiğimiz o değere kadar olan olasılıkların değerlerinin toplamı olan kümülatif fonksiyon. Bu da tabloda gördüğümüz değer.
* Aralık olarak verilseydi yine ikiye ayırıp küçük eşit şeklinde kullanmalıydık.
* Rare event az olan olay demekten ziyade terim gibi bir şey.
* ---
* In part b, instead of per day, our time period is now 2 days. So our expected value for average is 20 for 2 days. We changed the definition of our frequency. Instead of 1 day, we compute 2 days. 20 accounts per 2 days.
* Our is now 20.

All distributions are connected. You can move from one to another by changing the definition of random variable and do computation in other side.

MIDTERM IS UP TO HERE

Poisson Approximation of Binomial Distribution

Poisson distribution can be effectively used to approximate Binomial probabilities when the # of trials n is large, and the probability of success p is small.

* That can be considered as rare event.

Such approximation is adequate, say, for n>=30 and p<=0.05, and it becomes more accurate for larger n.

Example (New Accounts, continued – different question, same concept):

* Indeed, the situation in former example can be viewed as a sequence of Bernoulli trials.
* Suppose there are n = 400 000 potential internet users in the area, and on any specific day, each of them opens a new account with probability p = 0.000025.
* We see that the # of new accounts is the # of successes, hence a Binomial model with expectation E(X) = np = 10 is possible.
  + Only 10 of 400 000 people open an account.
* However, a distribution with such extreme n and p is unlikely to be found in any table, and computing its pmf by hand is tedious.
* Instead of Binomial distribution, one *can use Poisson distribution* with the same expectation which is = 10.
* ---
* Compute frequency and go to Poisson in short.
* Same event can be used for different distributions.

Text, letter

Description automatically generated

30 and 0.05 are made up numbers. You can say other numbers.

*Remark*: Mathematically, it means closeness of Binomial and Poisson pmf,

Text

Description automatically generated

and this is what S. D. Poisson has shown.

Number of trials is too big and probability is very small.

If you define np as and then using the limit, you can obtain the right hand side using the left hand side.

When p is large (p >= 0.95), the Poisson approximation is applicable too.

The probability of a failure “q = 1 - p” is small in this case.

Then, we can approximate the # of failures, which is also Binomial, by a Poisson distribution.

Example:

* 97% of electronic messages are transmitted with no error.
* What is the probability that out of 200 messages, at least 195 will be transmitted correctly?
* 