

This tutorial will make use of Python to help you gain a better understanding of families of discrete random variables. To import the required modules in Python, use

```
import numpy as np
from matplotlib import pyplot as plt
from scipy.stats import binom, geom, poisson, nbinom

#Example: Generate x values for plotting (adaptive range based on distribution)
x = np.arange(geom.ppf(0.01, p1)-2, geom.ppf(0.99, p1)+10).astype(int)
# Calculate PMF and CDF
pmf1 = geom.pmf(x, p1)
cdf1 = geom.cdf(x, p1)
```

Refer to the Scipy documentation to help you utilise the different distributions.

Use Google Colab to generate your code, a partial notebook can be found here: https://colab.research.google.com/drive/1AH2l_MIKqeBgPUGd-EX65VYpX5VE3n60?usp=sharing

- A factory's quality control inspector finds that, on average, 4 out of 5 products pass inspection. If 8 products are randomly selected for testing, what is the probability:

'n Fabriek se gehalte beheer inspekteur vind dat, gemiddeld, 4 uit 5 produkte die inspeksie slaag. As 8 produkte lukraak gekies word vir toetsing, wat is die waarskynlikheid:

 - that the factory passes more than 5 products? dat die fabriek meer as 5 produkte laat slaag?
 - that the factory fails more than once? dat die fabriek meer as een keer faal?
- A student takes practice quizzes for an exam. Each quiz has a probability $p = 0.35$ of being difficult, independent of previous quizzes. The student stops practicing once they encounter a difficult quiz. Let X denote the number of quizzes taken until encountering the first difficult one.

'n Student neem oefen toetse vir 'n eksamen. Elke toets het 'n waarskynlikheid van $p = 0.35$ om moeilik te wees, onafhanklik van vorige toetse. Die student stop oefen sodra hulle 'n moeilike toets teekom. Laat X die aantal toetse wees wat geneem is voor die eerste moeilike een teegekom word.

 - Plot the probability mass function $P_X(x)$ and cumulative distribution function $F_X(x)$ using Python (**Hint:** You may plot this in whichever way you wish, but it is suggested that you use `plt.vlines` to plot the PMF and `plt.step` (with `where="post"`) for the CDF). Plot die waarskynliheids massa funksie $P_X(x)$ en kumulatiewe verspreidingsfunksie $F_X(x)$ deur Python te gebruik (**Wenk:** Jy kan hierdie plot op enige manier, maar dit word aanbeveel om `plt.vlines` te gebruik vir die WMF en `plt.step` (met `where="post"`) vir die KVF).
 - Determine, by hand, the probability that the student will take more than 4 quizzes. Verify your answer using your Python plots. Bepaal, deur handberekeninge, die waarskynlikheid dat die student meer as 4 toetse sal neem. Verifieer jou antwoord deur jou Python plotte te gebruik.
 - The student switches to a different practice platform where quizzes are easier, with a probability of $p = 0.15$ of being difficult. The student still stops after encountering the first difficult quiz. Let X now denote the number of quizzes taken. Plot the new PMF and CDF of X in Python. Die student ruil na 'n ander oefen platform waar toetse makliker is, met 'n waarskynlikheid van $p = 0.15$ om moeilik te wees. Die student stop nogsteeds na die eerste moeilike toets. Laat X nou die aantal toetse wees wat geneem is. Plot die nuwe WMF en KVF van X in Python.
 - What is the probability that the student will take more than 4 quizzes on the new platform? Calculate this by hand and then use your Python plots to check your answer. Wat is die waarskynlikheid dat die student meer as 4 toetse op die nuwe platform sal neem? Bepaal hierdie met handberekeninge en gebruik dan jou plotte om seker te maak van jou antwoord.

3. When a server sends an email to a client, the probability that the email is delivered successfully is p . To ensure the message is received at least once, the server sends the same email n times. All transmissions are independent of one another.

(a) What is the probability mass function of K , the number of times the client receives the email in n transmissions?

(b) Assume that $p = 0.6$ and $n = 8$. Use Python to plot the PMF and CDF of K .

(c) How many times does the email need to be sent to produce a 98% probability that it is received correctly at least once? Calculate this by hand and then use the CDF plotted in Python to check your answer.

Wanneer 'n bediener 'n e-pos stuur na 'n klient, is die waarskynlikheid dat die e-pos suksesvol afgelewer word p . Om seker te wees dat die boodskap ten minste een keer ontvang word, stuur die bediener dieselfde e-pos n keer. Alle uitsendings is onafhanklik van mekaar.

Wat is die waarskynlikheids massa funksie van K , die hoeveelheid keer wat die klient die e-pos ontvang in n uitsendings?

Neem aan dat $p = 0.6$ en $n = 8$. Gebruik Python om die WMF en KVF van K te plot.

Hoeveel keer moet die e-pos gestuur word om 'n 98% waarskynlikheid te he dat dit ten minste een keer ontvang word? Gebruik handberekeninge om hierdie te bepaal en gebruik die KVF wat in Python geplot is om seker te maak van jou antwoord.
4. The number of customers that arrive at a coffee shop in T minutes is a Poisson random variable C with a rate of $\frac{1}{6}$ customers per minute.

(a) What is the probability that in a three-minute interval, exactly 2 customers will arrive? Calculate this by hand and check your answer by printing the PMF in Python.

(b) Plot the PMF $P_C(c)$ for a 9-minute interval using Python. Using this plot, determine the probability that exactly 3 customers will arrive in this time interval.

(c) How much time should be allowed so that the probability of at least one customers arriving is 0.95? Calculate this by hand and print the PMF out in Python to check your answer.

Die aantal kliënte wat binne T minute by 'n koffiewinkel aankom, is 'n Poisson-willekeurige veranderlike C met 'n tempo van $\frac{1}{6}$ kliënte per minuut.

Wat is die waarskynlikheid dat in 'n drie minute interval, presies 2 kliente sal arriveer? Gebruik handberekeninge en bepaal of jou antwoord reg is deur die WMF in Python uit te druk.

Plot die WMF $P_C(c)$ vir 'n 9-minuut interval deur gebruik te maak van Python. Gebruik hierdie plot om die waarskynlikheid te bepaal dat presies 3 kliente arriveer in hierdie interval.

Hoeveel tyd moet mens toelaat sodat die waarskynlikheid dat ten minste een kliente arriveer 0.95 is? Bepaal hierdie met handberekeninge en print die WMF uit in Python om seker te maak van jou antwoord.
5. A biotechnology company searches for breakthrough discoveries in their research labs. Historical data shows that each research project has a 12% chance of resulting in a breakthrough (**Tip:** For the Pascal distribution, you can use the `nbinom` class in Python as usual, just use `x+n` instead of `x` when using the `vlines` and `step` functions).

(a) Let X be the number of projects they must fund before achieving their n_{th} breakthrough. What is the PMF of X ?

(b) What is the probability that they achieve their first breakthrough on their 5th project? Calculate this by hand and check your answer by printing out the PMF in Python.

'n Biotegnologie maatskappy soek na deurbraak ontdekkings in hul navorsings laboratoriums. Historiese data wys dat elke navorsings projek 'n 12% kans het om 'n deurbraak te lewer (**Wenk:** Vir die Pascal verdeling, kan jy die `nbinom` klas in Python gebruik soos normaal, gebruik net `x+n` inplaas van `x` wanneer jy die `vlines` en `step` funksies gebruik).

Laat X die hoeveelheid projekte wees wat hulle moet befonds voor hulle hul n_{de} deurbraak behaal. Wat is die WMF van X ?

Wat is die waarskynlikheid dat hulle die eerste deurbraak op hul 5de projek behaal? Bereken eers hierdie self en maak seker van jou antwoord deur die WMF uit te druk in Python.

- (c) The company hopes to achieve 3 breakthroughs. They only have budget for 15 projects. You are tasked with finding out if this is feasible. Calculate by hand $P(X \leq 15)$. Now use Python to plot $P_X(x)$ and $F_X(x)$. Determine the expected value of X by hand and use Python to check your answer. Do you think they have a good chance to achieve those three breakthroughs?
- (d) A new study suggests that projects in a different research area have a 22% chance of breakthrough. The company considers shifting focus to this area, but it's more expensive, so they can only fund 12 projects to achieve their 3 breakthroughs. Like in the previous question, use hand calculations to determine $P(X \leq 12)$ and use Python to plot P_X and F_X . Use hand calculations to determine the new expected value of X . Should they shift their research focus?
- Die maatskappy hoop om 3 deurbrake te behaal. Hulle het net begroting vir 15 projekte. Jy moet uitvind of hierdie haalbaar is. Bereken met die hand $P(X \leq 15)$. Gebruik nou Python om P_X en F_X te plot. Bepaal die verwagte waarde van X met handberekeninge en gebruik Python om seker te maak van jou antwoord. Dink jy hulle het 'n goeie kans om daardie drie deurbrake te behaal?
- 'n Nuwe studie beweer dat projekte in 'n ander navorsings area 'n 22% kans het vir 'n deurbraak. Die maatskappy oorweeg om na hierdie area te skyf, maar dit is duurder, so hulle kan net 12 projekte befonds om hul 3 deurbrake te behaal. Soos in die vorige vraag, gebruik handberekeninge om $P(X \leq 12)$ te bereken en gebruik Python om P_X en F_X te plot. Bepaal die nuwe verwagte waarde van X met handberekeninge. Moet hulle hul navorsings fokus skyf?