Lecture 1

Poisson Distribution

The Poisson distribution is named after the French mathematician **Siméon Denis Poisson** (1781–1840) who introduced the distribution in 1837 in his work *Recherches sur la probabilité des jugements en matiére criminelle et en matiére civile* ("Research on the Probability of Judgments in Criminal and Civil Matters"). In his work the Poisson distribution describes the probability that a random event will occur in a time and/or space interval under the condition that the probability of any single event occurring is very small p, but the number of trials is very large N. So, for Siméon Denis Poisson, the Poisson(λ) distribution was a limit of binomial(p, N) distributions in the sense of the Law of Small Numbers 1.9: $p \to 0$ and $N \to \infty$, but $pN \to \lambda > 0$.

Another pioneer of the Poisson distribution was the Polish–German economist–statistician **Ladislaus Bortkiewicz** (1868–1931) who coined the term "Law of Small Numbers" in his 1898 investigation of the number of soldiers in the Prussian army killed accidentally by horse kick. Some have suggested that the Poisson distribution should be renamed the "Bortkiewicz distribution".

Siméon Denis Poisson (1781–1840)

The Poisson distribution is, in some sense, the uniform distribution on the natural numbers $\mathbb{N} = \{0, 1, 2, \ldots\}$. Indeed, the Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time and/or space if these events occur with a known average rate and independently. In the key example of this lecture, Example 1.1 below, the events are scattered in space, not in time.

1.1 Example (Malus Particles)
The Lake Diarrhea has, on average, 7 Malus particles per one liter. Magnus Flatus lives on the shore of the Lake Diarrhea. He drinks daily 2 liters of water from the Lake Diarrhea. The lethal daily intake of Malus particles is 30. What is the probability that Magnus Flatus will have a lethal intake of Malus particles in a given day?

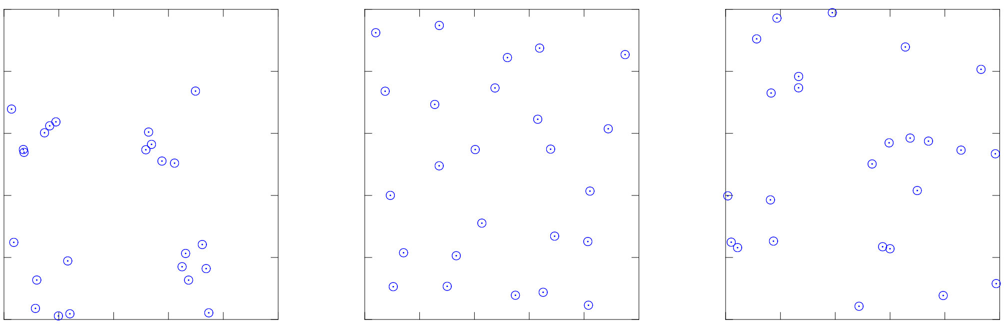
Qualitative Approach to Poisson Distribution

To answer the question of Example 1.1 we need to know the distribution of the random variable X that denotes the number of Malus particles in a 2 liter sample from the Lake Diarrhea. To fix the distribution of X we have to assume something about the distribution of the Malus particles in the lake. We know the average of the Malus particles: 7 per liter. Without any additional information, it is natural to assume that the particles are **independently and homogeneously scattered** in the lake. This means that knowledge of the amount of Malus particles in one sample does not help in predicting the amount of Malus particles in another sample (independence) and that samples taken from different parts of the lake are statistically the same (homogeneity). This leads us to the qualitative definition of the Poisson distribution, or actually the Poisson point process:

1.2 Definition (Poisson Point Process)
Let $\mathscr{A}$ be a collection of subsets of the Euclidean space $\mathbb{R}^d$ and let $\mathrm{vol}(A)$ denote the volume of the set A of $\mathbb{R}^d$. The family $X(A)$, $A \in \mathscr{A}$, is a **Poisson point process** with parameter $\lambda > 0$ if

(i) $X(A)$ takes values in $\mathbb{N} = \{0, 1, 2, \ldots\}$.
(ii) The distribution of $X(A)$ depends only on $\lambda\mathrm{vol}(A)$.
(iii) If A and B are disjoint, then $X(A)$ and $X(B)$ are independent.
(iv) $\mathbb{E}[X(A)] = \lambda\mathrm{vol}(A)$ for each A in $\mathscr{A}$.

In the context of Example 1.1 $\mathscr{A}$ a is the collection of water samples from the Lake Diarrhea and $X(A)$ is the number of Malus particles in the water sample A. The Lake Diarrhea itself is a subset of $\mathbb{R}^3$.



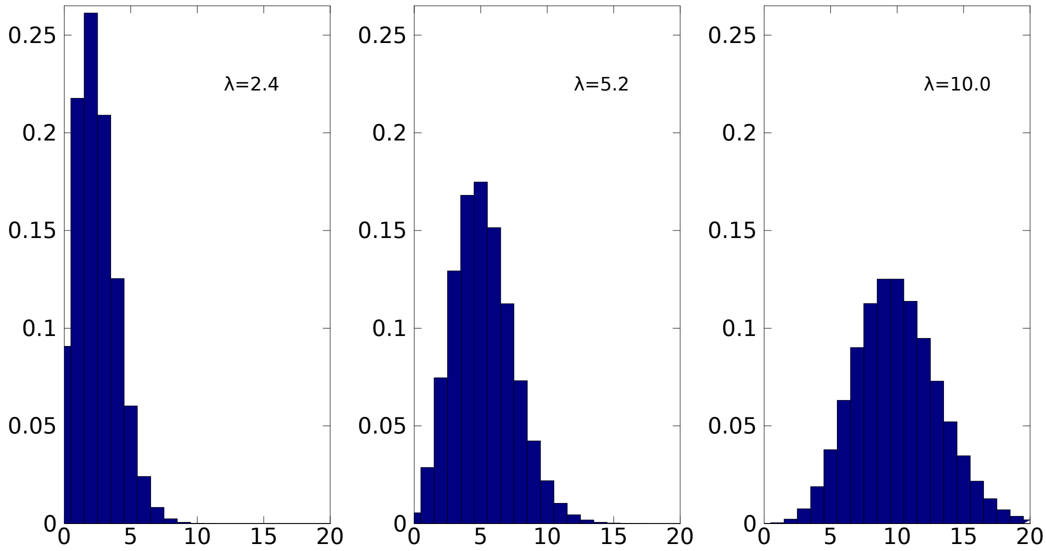
Samples of point processes. See `point_processes.m` to see which one is the Poisson point process.

Quantitative Approach to Poisson Distribution

The qualitative definition 1.2 does not, yet, allow calculations of probabilities, although some expectations can be calculated. Indeed, we already did calculate some expectations in Exercise 1.4. However, it turns out that the qualitative definition 1.2 actually fixes the distributions completely as Theorem 1.4 will show. Before that, let us recall the Poisson distribution.

1.3 Definition (Poisson Distribution)
A random variable X has the **Poisson distribution** with parameter $\lambda > 0$ if it has the probability mass function

$$\mathbb{P}[X = x] \quad = \quad e^{-\lambda}\frac{\lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots.$$



Probability mass functions of Poisson distribution with different parameters λ.

The plots above were generated with Octave with the following code:

```
##########################################################################
## FILE: poisson_pmfs.m
##
## Plots some probability mass functions (pmf) of Poisson distributions.
##########################################################################

## Data for plots.
lambda = [2.4, 5.2, 10.0];          # Intensities for the plots.
x = 0:20;                           # The x's for Px=prob(x) to be plotted.
w = 1;                              # Width of the bar in the bar plot.
Px(1,:) = poisspdf(x, lambda(1));   # 1st row for lambda(1).
Px(2,:) = poisspdf(x, lambda(2));   # 2nd row for lambda(2).
Px(3,:) = poisspdf(x, lambda(3));   # 3rd row for lambda(3).

```

```
## Plotting (bar plots).
plotlims = [0, 20, 0, 0.265];          # Plotting window [x1, x2, y1, y2].
subplot(1,3,1);                        # 1 row, 3 columns, 1st plot.
    bar(x, Px(1,:), w);
    text(12, 0.225, '\lambda=2.4');
    axis(plotlims);
subplot(1,3,2);                        # 1 row, 3 columns, 2nd plot.
    bar(x, Px(2,:), w);
    text(12, 0.225, '\lambda=5.2');
    axis(plotlims);
subplot(1,3,3);                        # 1 row, 3 columns, 3rd plot.
    bar(x, Px(3,:), w);
    text(12, 0.225, '\lambda=10.0');
    axis(plotlims);

## Crop the borders of the current figure and print landscape oriented PDF.
orient('landscape');
papersize = get(gcf, 'papersize');
papersize(2) = 0.75*papersize(2);
set(gcf, 'papersize', papersize);
border = -0.75;
set(gcf, 'paperposition', [border, 0, papersize(1)-2*border, papersize(2)]);
print('poisson_pmfs.pdf');
```

www.uva.fi/~tsottine/sp_with_octave/poisson_pmfs.m

1.4 Theorem (Poisson Point Process is Poisson Distributed)
For the Poisson point process $X(A)$, $A \in \mathscr{A}$, of Definition 1.2 it must hold true that

$$\mathbb{P}\left[X(A)=x\right] \quad = \quad e^{-\lambda \mathrm{vol}(A)} \frac{(\lambda \mathrm{vol}(A))^x}{x!} \qquad \text{for all } A \text{ in } \mathscr{A} \text{ and } x = 0, 1, \ldots,$$

where $\mathrm{vol}(A)$ is the volume of the set A.

Let us argue how the Poisson distribution arises from the Poisson point process, i.e, let us argue why Theorem 1.4 holds. Denote

$$p\big(x; \mathrm{vol}(A)\big) \quad = \quad \mathbb{P}\left[X(A)=x\right].$$

The key idea in the argument is the following: Split A into two parts with volumes v and w. Then, because of independence and homogeneity, we obtain the functional equation

$$p(x; v+w) \quad = \quad \sum_y p(x-y; v)p(y; w).$$

We argue by using probability generating functions that this functional equation implies that $p(x; v)$ is a Poisson distribution.

Recall that the **probability generating function** of an $\mathbb{N}$-valued random variable X is

$$G(\theta) \quad = \quad \sum_x \mathbb{P}[X=x]\theta^x.$$

Moreover, recall the following facts of the probability generating functions that we state here as a lemma.

1.5 Lemma (Properties of Probability Generating Functions)
The probability generating function $G(\theta)$ of an $\mathbb{N}$-valued random variable X determines the distribution of X. In particular:

(i) $G^{(x)}(0) = x!\,\mathbb{P}[X = x]$
(ii) $G(1) = 1$
(iii) $G'(1) = \mathbb{E}[X]$
(iv) $G''(1) + G'(1) - G'(1)^2 = \mathbb{V}\mathrm{ar}[X]$

The proof of Lemma 1.5 is left as an exercise.

Let $G(\theta; v + w)$ be the probability generating function of the random variable $X(A)$. Then, by the functional equation coming from the splitting argument we obtain that

$$\begin{aligned}
G(\theta; v+w) &= \sum_x p(x; v+w)\theta^x \\
&= \sum_x \sum_y p(x-y; v)p(y; w)\theta^x \\
&= \sum_y \sum_x p(x-y; v)\theta^{x-y} p(y; w)\theta^y \\
&= \sum_z \sum_x p(z; v)\theta^z p(y; w)\theta^y \\
&= G(\theta; v)G(\theta; w).
\end{aligned}$$

Denote $g(\theta; v) = \log G(\theta; v)$. Then from the above we obtain the **Cauchy's functional equation**

$$g(\theta; v+w) = g(\theta; v) + g(\theta; w).$$

So, $g(\theta; v)$ is **additive** in v. Since $g(\theta; v)$ is also increasing in v, it follows that $g(\theta; v) = v\psi(\theta)$ for some $\psi(\theta)$. So $G(\theta; v) = e^{v\psi(\theta)}$. Since $G(\theta; v)$ is a probability generating function we must have, by Exercise 1.3

$$\begin{aligned}
G(1; v) &= 1, \\
G'(1; v) &= \mathbb{E}[X(A)] = \lambda v.
\end{aligned}$$

Thus, we must have $\psi(\theta) = \lambda(\theta - 1)$. So,

$$G(\theta; v) = e^{\lambda v(\theta-1)}.$$

Since probability generating functions determine probabilities, the claim follows from the following exercise.

1.6 Example (Malus Particles, Solution)
Now we know that the number of Malus particles Magnus Flatus consumes daily has the distribution

$$\mathbb{P}[X=x] \quad = \quad e^{-14}\frac{e^{14x}}{x!}.$$

The probability in question is

$$\begin{aligned}\mathbb{P}[X\geq 30] &= 1-\mathbb{P}[X\leq 29] \\ &= 1-\sum_{x=0}^{29} e^{-14}\frac{e^{14x}}{x!}.\end{aligned}$$

Since we do not want to calculate all the 30 terms of the sum above by hand, we use Octave. The simplest way of doing this is to call the Octave function `poisscdf`. So, typing `1-poisscdf(29,14)` we get the answer 0.01358 %.

Sums of Independent Poisson Distributions

From the qualitative approach to the Poisson process it is **intuitively** clear that if X_1 and X_2 are **independent** Poisson distributed with parameters λ_1 and λ_2, respectively, then their sum X_1+X_2 is also Poisson distributed with parameter $\lambda_1+\lambda_2$. **Rigorously** this can be seen from the following calculations:

$$\begin{aligned}\mathbb{P}\left[X_1+X_2=x\right] &= \sum_{x_1=0}^{\infty}\mathbb{P}\left[X_1=x_1, X_2=x-x_1\right] \\ &= \sum_{x_1=0}^{x}\mathbb{P}\left[X_1=x_1\right]\mathbb{P}\left[X_2=x-x_1\right] \\ &= \sum_{x_1=0}^{x} e^{-\lambda_1}\frac{\lambda_1^{x_1}}{x_1!}e^{-\lambda_2}\frac{\lambda_2^{x-x_1}}{(x-x_1)!} \\ &= e^{-(\lambda_1+\lambda_2)}\sum_{x_1=0}^{x}\frac{\lambda_1^{x_1}}{x_1!}\frac{\lambda_2^{x-x_1}}{(x-x_1)!} \\ &= e^{-(\lambda_1+\lambda_2)}\frac{1}{x!}\sum_{x_1=0}^{x}\frac{x!}{x_1!(x-x_1)!}\lambda_1^{x_1}\lambda_2^{x-x_1} \\ &= e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^x}{x!}.\end{aligned}$$

Here the last equality followed from the **binomial formula**

$$\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k} \quad = \quad (a+b)^n.$$

Repeating the arguments above for n summands, we obtain the following:

1.7 Proposition (Poisson Sum)
Let $X_1, X_2, \ldots, X_n$ be independent Poisson distributed random variables with parameters $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively. Then their sum $X_1 + X_2 + \cdots + X_n$ is Poisson distributed with parameter $\lambda_1 + \lambda_2 + \cdots + \lambda_n$.

Let us then consider a reverse of Proposition 1.7. Suppose X_1 and X_2 are **independent** Poisson distributed random variables with parameters λ_1 and λ_2, respectively. Suppose further that we know the value of their sum: $X_1 + X_2 = x$. What can we say about X_1? **Intuitively**, we can argue as follows: each point of the Poisson point process $X_1 + X_2$ comes independently from either X_1 or X_2. The relative contribution of X_1 to the points is $\lambda_1/(\lambda_1+\lambda_2)$. So, this is the probability of success, if success means that the point comes from the random variable X_1. Since these successes are independent we arrive at the binomial distribution: X_1 is binomially distributed with parameters x and $\lambda_1/(\lambda_1+\lambda_2)$. **Rigorously**, the educated guess above is seen to be true from the following calculations: First, simple use of definitions yield

$$\begin{aligned}
\mathbb{P}\left[X_1 = x_1 \,|\, X_1 + X_2 = x\right] &= \frac{\mathbb{P}\left[X_1 = x_1, X_2 = x - x_1\right]}{\mathbb{P}\left[X_1 + X_2 = x\right]} \\
&= \frac{\mathbb{P}\left[X_1 = x_1\right]\mathbb{P}\left[X_2 = x - x_1\right]}{\mathbb{P}\left[X_1 + X_2 = x\right]} \\
&= e^{-\lambda_1}\frac{\lambda_1^{x_1}}{x_1!}e^{-\lambda_2}\frac{\lambda_2^{x-x_1}}{(x-x_1)!}\bigg/ e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^x}{x!}
\end{aligned}$$

Then rearranging the terms in the result above yields

$$\mathbb{P}\left[X_1 = x_1 \,|\, X_1 + X_2 = x\right] = \binom{x}{x_1}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{x_1}\left(1 - \frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{x-x_1}.$$

So, we see that X_1 given $X_1 + X_2 = x$ is binomially distributed with parameters x and $\lambda_1/(\lambda_1+\lambda_2)$. Repeating the arguments above for n summands, we obtain the following:

1.8 Proposition (Reverse Poisson Sum)
Let $X_1, X_2, \ldots, X_n$ be independent Poisson distributed random variables with parameters $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively. Let $X = X_1 + X_2 + \cdots + X_n$ and $\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_n$. Then, conditionally on $X = x$ the random variables X_k, $k = 1, 2, \ldots, n$, are binomially distributed with parameters x and λ_k/λ.

Proposition 1.8 gave one important connection between the Poisson and the binomial distributions. There is another important connection between these distributions. This connection, the Law of Small Numbers 1.9 below, is why Siméon Denis Poisson introduced the Poisson distribution.

1.9 Theorem (Law of Small Numbers)
Let X_n, $n \in \mathbb{N}$, be binomially distributed with parameters n and p_n. Suppose that $np_n \to \lambda$ as $n \to \infty$. Then the distribution of X_n converges to the Poisson distribution with parameter λ, i.e.,

$$\binom{n}{x} p_n^x (1-p_n)^{n-x} \quad \to \quad e^{-\lambda}\frac{\lambda^x}{x!} \qquad \text{for all } x = 0, 1, 2, \ldots,$$

whenever $np_n \to \lambda$.

The proof of Theorem 1.9 is Exercise 1.13. Maybe the easiest way to prove Theorem 1.9 is to use the **Stirling's formula** for the factorial:

$$n! \;\sim\; \sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n.$$

Here the asymptotic notation $a_n \sim b_n$ means that

$$\lim_{n\to\infty} \frac{a_n}{b_n} \;=\; 1.$$

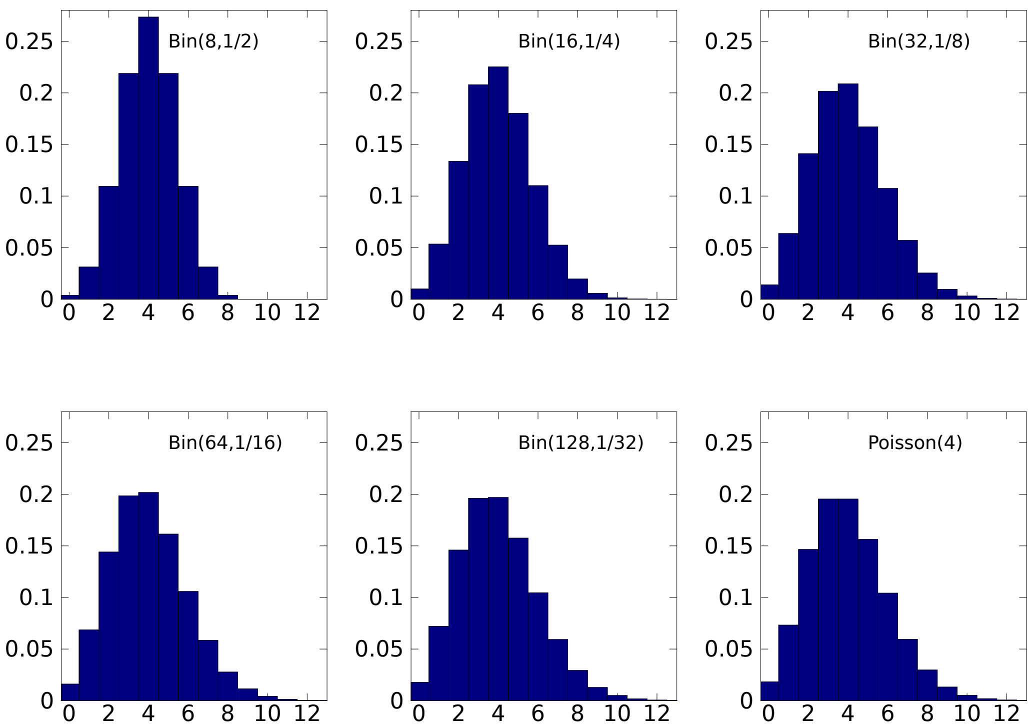


Illustration of how the Binomial becomes Poisson by the Law of Small Numbers.

The plots above were generated with Octave with the following code:

```
###########################################################################
## FILE: poisson_binomial.m
##
## An illustration of the Law of Small Numbers, i.e., how binomial
## distributions approximate the Poisson distribution, or vice versa.
##
## This is quick and very dirty coding. No-one should learn this style!
###########################################################################


lambda = 4;                       # Limit lambda=p*n is fixed.
w = 1;                            # The width of the column for the bar plot.
plotlims = [-0.4, 13, 0, 0.28];   # Plotting window [x1, x2, y1, y2].

## The 2x3 subplots.
x = 0:8;
n = 8;
P = binopdf(x,n,lambda/n);
subplot(2,3,1)
    bar(x, P, w);
    text(5, 0.25, 'Bin(8,1/2)');
    axis(plotlims);

x=0:15;
n = 16;
P = binopdf(x,n,lambda/n);
subplot(2,3,2)
    bar(x, P, w);
    text(5, 0.25, 'Bin(16,1/4)');
    axis(plotlims);

x = 0:15;
n = 32;
P = binopdf(x,n,lambda/n);
subplot(2,3,3)
    bar(x, P, w);
    text(5, 0.25, 'Bin(32,1/8)');
    axis(plotlims);

x = 0:15;
n = 64;
P = binopdf(x,n,lambda/n);
subplot(2,3,4)
    bar(x, P, w);
    text(5, 0.25, 'Bin(64,1/16)');
    axis(plotlims);

x = 0:15;
n = 256;
P = binopdf(x,n,lambda/n);
subplot(2,3,5)
    bar(x, P, w);
    text(5, 0.25, 'Bin(128,1/32)');
    axis(plotlims);
```

```

x= 0:15;
P = poisspdf(x,lambda);
subplot(2,3,6);
    bar(x, P, w);
    text(5, 0.25, 'Poisson(4)');
    axis(plotlims);


## Crop the borders of the current figure and print landscape oriented PDF.
orient('landscape');
papersize = get(gcf, 'papersize');
#papersize(2) = 0.75*papersize(2);  # I ma not sure if this works as it should!
set(gcf, 'papersize', papersize);
border = -0.75;
set(gcf, 'paperposition', [border, 0, papersize(1)-2*border, papersize(2)]);
print('poisson_binomial.pdf');
```

www.uva.fi/~tsottine/sp_with_octave/poisson_binomial.m

Exercises

In the exercises "calculate" has the 21^{st} century meaning. I.e., you can calculate with pen and paper, with pocket calculator, with Excel, or with Octave (recommended). So, in the modern tongue "calculate" means "program".

1.1 Exercise
Let X be Poisson distributed with parameter 1.5. Calculate

(a) $\mathbb{P}[X = 1.5]$
(b) $\mathbb{P}[X = 0]$
(c) $\mathbb{P}[X = 0 \text{ or } X = 10]$
(d) $\mathbb{P}[X < 1.5]$
(e) $\mathbb{P}[X > 1.5]$,
(f) $\mathbb{P}[1 < X \leq 10]$

1.2 Exercise
The probability generating function is named thus because the probabilities $\mathbb{P}[X = x]$ of an $\mathbb{N}$-valued random variable X can be recovered from the probability generating function $G(\theta)$ of X by differentiating:

$$\mathbb{P}[X = x] \quad = \quad \frac{1}{x!} G^{(x)}(0), \tag{1.10}$$

where $G^{(x)}(0)$ is the x^{th} derivative of the function $G(\theta)$ evaluated at point $\theta = 0$.

Prove the formula (1.10), and conclude from it that the probability generating function determines the distribution of an $\mathbb{N}$-valued random variable uniquely.

1.3 Exercise
Let $G(\theta)$ be the moment generating function of and $\mathbb{N}$-valued random variable X. Show that

(a) $1 = G(1)$

(b) $\mathbb{E}[X] = G'(1)$

(c) $\mathbb{E}[X(X-1)] = G''(1)$

(d) $\mathbb{V}\mathrm{ar}[X] = G''(1) + G'(1) - G'(1)^2$

1.4 Exercise
Consider Example 1.1.

(a) How many Malus particles, on average, would Magnus Flatus get daily if he would drink only 1 liter per day?

(b) Suppose Magnus Flatus wants to get, on average, only 10 Malus particles per day. How many liters can he drink from the Lake Diarrhea daily?

(c) Suppose Magnus Flatus wants to get, on average, only 30 Malus particles per year. How many liters can he drink from the Lake Diarrhea daily?

1.5 Exercise
Show that the definition for the probability mass function of the Poisson distribution in Definition 1.3 is correct, i.e., it is non-negative and sums up to one.

1.6 Exercise
Let X be Poisson distributed with parameter λ. Show that

(a) $\mathbb{E}[X] = \lambda$

(b) $\mathbb{V}\mathrm{ar}[X] = \lambda$.

1.7 Exercise
Show that the probability generating function of a Poisson distribution with parameter λ is $G(\theta) = e^{\lambda(\theta-1)}$.

1.8 Exercise
Consider Magnus Flatus from Example 1.1.

(a) How much can Magnus Flatus drink daily from the Lake Diarrhea so that his probability of daily overdose of 30 particles is still under $1\,\%$?

(b) Magnus Flatus drinks 2 liters from the Lake Diarrhea for 40 years. What is the probability that his daily intake of Malus particles will never exceed the lethal level of 30 particles?

(c) How many liters of water from the Lake Diarrhea can Magnus Flatus drink daily for 40 years so that the probability of a lethal daily intake is still less than $1\,\%$?

1.9 Exercise

Let X_1, X_2 and X_3 be Poisson distributed with parameters 2.4, 5.2 and 10.0, respectively. Calculate the probabilities

(a) $\mathbb{P}[X_1 + X_2 = 0]$

(b) $\mathbb{P}[X_1 + X_2 = 1]$

(c) $\mathbb{P}[X_1 + X_2 + X_3 = 5]$

(d) $\mathbb{P}[X_1 + X_2 \leq 10]$

(e) $\mathbb{P}[1 \leq X_1 + X_2 \leq 10]$

(f) $\mathbb{P}[5 < X_1 + X_2 + X_3 \leq 10]$

1.10 Exercise

Let X_1, X_2 and X_3 be Poisson distributed with parameters 2.4, 5.2 and 10.0, respectively. Calculate the conditional probabilities

(a) $\mathbb{P}[X_1 = 0 \,|\, X_1 + X_2 + X_3 = 1]$

(b) $\mathbb{P}[X_1 = 3 \,|\, X_2 + X_2 + X_3 = 3]$

(c) $\mathbb{P}[X_3 = 1 \,|\, X_1 + X_2 + X_3 = 2]$

(d) $\mathbb{P}[X_1 \leq 1 \,|\, X_1 + X_2 + X_3 = 1]$

(e) $\mathbb{P}[2 \leq X_1 \leq 3 \,|\, X_2 + X_2 + X_3 = 3]$

(f) $\mathbb{P}[X_3 > 0 \,|\, X_1 + X_2 + X_3 = 10]$

1.11 Exercise

Let X_1 be Poisson distributed with parameter 2 and let X_2 be Poisson distributed with parameter 5. Suppose $X_1 + X_2 = 10$. What is the probability that $X_1 > X_2$?

1.12 Exercise

Consider Magnus Flatus from Example 1.1. Magnus Flatus has sampled 10 liters from the Lake Diarrhea. He analyzed the sample and found 50 Malus particles. Nevertheless, Magnus Flatus bottles the 10 liter sample into 20 half-liter bottles and takes one of the bottles with him when he goes hiking in the next weekend. What is the probability that the bottle contains a lethal dose of Malus particles?

1.13 Exercise ⋆
Explain why Law of Small Numbers 1.9 is named thus and prove the law.