Universität Bielefeld
Tanja Pasurek
Ehsan Abedi WiSe 2023/24

Foundations of Statistics

Solutions to Homework 5

Lecture material: Chapters 1.4–1.6

Exercise 1. (Normal distribution). Let X be a $\mathcal{N}(\mu, \sigma^2)$ distributed random variable with probability density function (PDF) $\phi_{\mu,\sigma^2}(x)$ and distribution function (CDF) $\Phi_{\mu,\sigma^2}(x) := \mathbb{P}(X \leq x)$.

a Show that the distribution of the standardized random variable $Z := \frac{X-\mu}{\sigma}$ is $\mathcal{N}(0,1)$ (= *standard normal distribution*).

Solution: We have

$$\begin{aligned} f_X(x) = \phi_{\mu,\sigma^2}(x) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right], \\ F_X(x) = \Phi_{\mu,\sigma^2}(x) &= \int_{-\infty}^{x} \phi_{\mu,\sigma^2}(y)\,\mathrm{d}y = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right] \mathrm{d}y. \end{aligned}$$

We calculate the CDF of the random variable Z

$$\begin{aligned} F_Z(x) &:= \mathbb{P}(Z \leq x) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \leq x\right) = \mathbb{P}\left(X \leq \mu + \sigma x\right) \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\mu+\sigma x} \exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right] \mathrm{d}y \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mu+\sigma x} \exp\left[-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right] \mathrm{d}\left(\frac{y-\mu}{\sigma}\right) \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left[-z^2/2\right] \mathrm{d}z = \Phi_{0,1}(x), \end{aligned}$$

where we introduced a new variable $z := \frac{y-\mu}{\sigma}$. Note that the upper integration limit $y = \mu + \sigma x$ corresponds to $z = \frac{\mu+\sigma x-\mu}{\sigma} = x$. Hence, $Z \sim N(0,1)$.

b Show $\mathbb{P}(X \leq b) = \Phi_{0,1}\left(\frac{b-\mu}{\sigma}\right)$ and deduce a formula for $\mathbb{P}(a \leq X \leq b)$.

Solution: Using part (a)

$$\mathbb{P}(X \leq b) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) = \mathbb{P}\left(Z \leq \frac{b-\mu}{\sigma}\right) = \Phi_{0,1}\left(\frac{b-\mu}{\sigma}\right).$$

To find $\mathbb{P}(a \leq X \leq b)$, first note that

$$\{X < a\} \cup \{a \leq X \leq b\} = \{X \leq b\}$$

then observe that

$$\begin{aligned} \mathbb{P}(a \leq X \leq b) &= \mathbb{P}(X \leq b) - \mathbb{P}(X < a) \\ &= \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a) \qquad (X \text{ is a continuous r.v.}) \\ &= \Phi_{0,1}\left(\frac{b-\mu}{\sigma}\right) - \Phi_{0,1}\left(\frac{a-\mu}{\sigma}\right), \end{aligned}$$

c Show $\Phi_{0,1}(x) + \Phi_{0,1}(-x) = 1$.

Solution: By definition of the CDF

$$\Phi_{0,1}(-x) = \int_{-\infty}^{-x} \phi_{0,1}(y)\,\mathrm{d}y = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-x} \exp\left(-y^2/2\right)\,\mathrm{d}y = \frac{1}{\sqrt{2\pi}}\int_{x}^{+\infty} \exp\left(-z^2/2\right)\,\mathrm{d}z,$$

where we used the symmetry of $\phi_{0,1}(y)$ and the change of variable $z = -y$. Hence

$$\begin{aligned} \Phi_{0,1}(x) + \Phi_{0,1}(-x) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \exp\left(-y^2/2\right)\,\mathrm{d}y + \frac{1}{\sqrt{2\pi}}\int_{x}^{+\infty} \exp\left(-z^2/2\right)\,\mathrm{d}z \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \exp\left(-y^2/2\right)\,\mathrm{d}y = 1. \end{aligned}$$

d For $\mu = 0$ and $\sigma^2 = 1$ find b with `R` such that $\mathbb{P}(-b \leq X \leq b) = 0.8$.

Solution: By the symmetry

$$\mathbb{P}(X > b) = \mathbb{P}(X < -b) = \frac{1}{2}(1 - 0.8) = 0.1,$$
$$\mathbb{P}(X \leq b) = 1 - \mathbb{P}(X > b) = 0.9.$$

```
b=qnorm(0.9, mean=0, sd=1)
```

```
[1] 1.281552
```

or

```
b=qnorm(0.9)
```

```
[1] 1.281552
```

e Use parts b and c to show that the value of $\mathbb{P}(\mu - \sigma \leq X \leq \mu + \sigma)$ does not depend on μ, σ. Calculate its value in R.

Solution: In part b, we showed that

$$\mathbb{P}(a \leq X \leq b) = \Phi_{0,1}\left(\frac{b-\mu}{\sigma}\right) - \Phi_{0,1}\left(\frac{a-\mu}{\sigma}\right).$$

Here in particular, we take $a = \mu - \sigma$ and $b = \mu + \sigma$. Then

$$\begin{aligned}
\mathbb{P}(\mu - \sigma \leq X \leq \mu + \sigma) &= \Phi_{0,1}\left(\frac{\mu + \sigma - \mu}{\sigma}\right) - \Phi_{0,1}\left(\frac{\mu - \sigma - \mu}{\sigma}\right) \\
&= \Phi_{0,1}(1) - \Phi_{0,1}(-1) \\
&= \{1 - \Phi_{0,1}(-1)\} - \Phi_{0,1}(-1) \qquad \text{(part c)} \\
&= 1 - 2\Phi_{0,1}(-1)
\end{aligned}$$

so the result does not depend on concrete values of μ and σ.

```
1-2*pnorm (-1)
```

```
[1] 0.6826895
```

f Generate 10000 random samples of X with arbitrary numeric values μ and σ and verify the result of (e) with a suitable simulation in R.

Solution:

```
m <- -4; s <- 3; # try different values for mean and var.
n <- 10000
vec <- rnorm(n, mean=m, sd=s)
sum((vec >= m - s) & (vec <= m + s))/n
```

```
[1] 0.6883
```

Note that this is just an approximation and we would expect that as $n \to \infty$, this value converges to the result of part e.

Exercise 2. (Sum of two independent random variables). The goal of this exercise is to study the distribution of sum of two independent random variables.

a Let X, Y be two independent discrete random variables with PMF f_X and f_Y. Prove that the PMF of $Z = X + Y$ is given by

$$f_Z(z) = \sum_u f_X(z-u) f_Y(u). \tag{1}$$

Solution: The PMF of Z can be written as

$$\begin{aligned}
f_Z(z) :=& \mathbb{P}(Z = z) \qquad \Big(= \mathbb{P}(\{\omega \in \Omega : Z(\omega) = z\})\Big) \\
=& \mathbb{P}(X + Y = z) \\
=& \sum_u \mathbb{P}(\{X + Y = z\} \cap \{Y = u\}) \qquad \text{(Law of total probability)} \\
=& \sum_u \mathbb{P}(\{X = z - u\} \cap \{Y = u\}) \\
=& \sum_u \mathbb{P}(X = z - u)\mathbb{P}(Y = u) \qquad \text{(Independence)} \\
=& \sum_u f_X(z-u) f_Y(u)
\end{aligned}$$

where obviously the sum runs over all $u \in \mathbb{R}$ such that $f_X(z-u) \neq 0$ and $f_Y(u) \neq 0$, otherwise the summands are zero.

b Now let X, Y be two independent continuous random variables with PDF f_X and f_Y. Prove that the PDF of $Z = X + Y$ is given by

$$f_Z(z) = \int_{u=-\infty}^{u=\infty} f_X(z-u) f_Y(u)\, \mathrm{d}u. \tag{2}$$

which is the convolution of their respective PDFs.
Hint: First derive the CDF of Z,

$$F_Z(z) := \mathbb{P}(Z \leq z).$$

Solution: Let us first define the following set for a given $z \in \mathbb{R}$:

$$A(z) := \big\{(x, y) \in \mathbb{R} \times \mathbb{R} : \quad x + y \leq z\big\}.$$

Now, we derive the CDF of Z:

$$
\begin{aligned}
F_Z(z) :=& \mathbb{P}(Z \leq z) \qquad \Big(= \mathbb{P}(\{\omega \in \Omega : Z(\omega) \leq z\})\Big) \\
=& \mathbb{P}(X + Y \leq z) \\
=& \int_{A(z)} f_{(X,Y)}(x, y) dx dy \\
=& \int_{A(z)} f_X(x) f_Y(y) dx dy \qquad\qquad \text{(independence)} \\
=& \int_{\mathbb{R}\times\mathbb{R}} \mathbb{1}_{A(z)}(x, y) f_X(x) f_Y(y) dx dy \\
=& \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbb{1}_{A(z)}(x, y) f_X(x) dx \right) f_Y(y) dy \\
=& \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z-y} f_X(x) dx \right) f_Y(y) dy \\
=& \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z} f_X(x - y) dx \right) f_Y(y) dy \\
=& \int_{-\infty}^{z} \Bigg(\underbrace{\int_{-\infty}^{\infty} f_X(x - y) f_Y(y) dy}_{(f_X * f_Y)(x)} \Bigg) dx
\end{aligned}
$$

This concludes the proof that $f_Z = f_X * f_Y$.

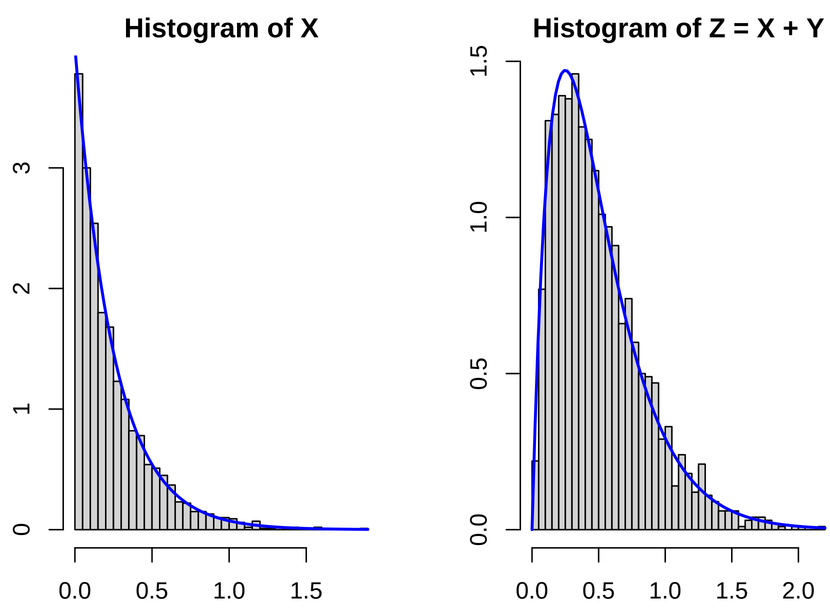
c Use formula (2) to find how $Z = X + Y$ is distributed if $X \sim Exp(\lambda)$ and $Y \sim Exp(\lambda)$ are independent. To illustrate the result, pick some particular $\lambda > 0$. Use `rexp()` in `R` to generate random samples. Create two plots: one with histogram of samples of X and density function of exponential distribution, and the other with histogram of samples Z and the density function that you have found.

Solution: We first note that $Z \in [0, \infty)$. For $z \geq 0$, we have

$$
\begin{aligned}
f_Z(z) &= \int_{-\infty}^{\infty} f_X(z - u) f_Y(u) \, \mathrm{d}u \\
&= \int_0^z \lambda e^{-\lambda(z-u)} \lambda e^{-\lambda(u)} \, \mathrm{d}u \\
&= \lambda^2 \int_0^z e^{-\lambda(z)} \, \mathrm{d}u = \lambda^2 z e^{-\lambda z},
\end{aligned}
$$

which is *Gamma* distribution with shape 2 and rate parameter λ (in this case also known as *Erlang* distribution with shape 2 and rate λ.)

```r
# Take an arbitrary parameter
set.seed(3)
lambda <- 4

X <- rexp(2000, rate = lambda)
Y <- rexp(2000, rate = lambda)
Z <- X + Y

# theoretical PDF of Z
pdf_Z <- function(z, lambda) {
  lambda^2 * z * exp(-lambda * z)
}

par(mfrow = c(1, 2), mar = c(3, 3, 2, 2))

# histogram of the samples of X
hist(X, breaks = 40, freq = FALSE, main = "Histogram of X",
     xlab = "x", ylab = "Density")
curve(dexp(x,lambda), col = "blue", lwd = 2, add = TRUE)

# histogram of the samples of Z = X + Y
hist(Z, breaks = 40, freq = FALSE, main = "Histogram of Z = X + Y",
     xlab = "z", ylab = "Density")
curve(pdf_Z(x, lambda), col = "blue", lwd = 2, add = TRUE)
```



d (optional*) Use formula (2) to find how $Z = X + Y$ is distributed if $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$ are independent and normally distributed. Compare the result with computation of $\mathbb{E}(Z)$ and $\mathrm{Var}(Z)$.

Solution: In this exercise, we use the notation $f_{\mu,\sigma^2}(x)$ to denote density function corresponding to the normal distribution $\mathcal{N}(\mu, \sigma^2)$. We have

$$
\begin{aligned}
f_X(x) &= \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left[-\frac{1}{2\sigma_1^2}(x-\mu_1)^2\right] =: f_{\mu_1,\sigma_1^2}(x), \\
f_Y(x) &= \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left[-\frac{1}{2\sigma_2^2}(x-\mu_2)^2\right] =: f_{\mu_2,\sigma_2^2}(x).
\end{aligned}
$$

Therefore we obtain

$$
\begin{aligned}
f_Z(x) &= f_X * f_Y(x) \\
&= \int_{-\infty}^{\infty} f_X(x-y) f_Y(y) dy \\
&= \int_{-\infty}^{\infty} f_X(y) f_Y(x-y) dy \\
&= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(y-\mu_1)^2}{2\sigma_1^2}\right) \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(x-y-\mu_2)^2}{2\sigma_2^2}\right) dy \\
&= \frac{1}{2\pi\sigma_1\sigma_2} \int_{\mathbb{R}} \exp\left(-\frac{(y-\mu_1)^2}{2\sigma_1^2}\right) \exp\left(-\frac{(x-y-\mu_2)^2}{2\sigma_2^2}\right) dy.
\end{aligned}
$$

Next, we use the linear transformation $y \mapsto y + \mu_1$:

$$
\begin{aligned}
&f_Z(x) \\
=&\frac{1}{2\pi\sigma_1\sigma_2} \int_{\mathbb{R}} \exp\left(-\frac{y^2}{2\sigma_1^2}\right) \exp\left(-\frac{(x-y-\mu_1-\mu_2)^2}{2\sigma_2^2}\right) dy \\
=&\frac{1}{2\pi\sigma_1\sigma_2} \int_{\mathbb{R}} \exp\left(-\frac{\sigma_2^2 y^2 + \sigma_1^2(x-y-\mu_1-\mu_2)^2}{2\sigma_1^2\sigma_2^2}\right) dy \\
=&\frac{1}{2\pi\sigma_1\sigma_2} \int_{\mathbb{R}} \exp\left(-\frac{\left(y\sqrt{\sigma_1^2+\sigma_2^2} - (x-\mu_1-\mu_2)\frac{\sigma_1^2}{\sqrt{\sigma_1^2+\sigma_2^2}}\right)^2}{2\sigma_1^2\sigma_2^2} + \frac{(x-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}\right) dy.
\end{aligned}
$$

Now we can pull out the factor below as it does not depend on y

$$
\exp\left(-\frac{(x-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}\right)
$$

and recall the formula of $f_{\mu_1+\mu_2,\sigma_1^2+\sigma_2^2}(x)$:

$$\begin{aligned}
&f_Z(x)\\
=&f_{\mu_1+\mu_2,\sigma_1^2+\sigma_2^2}(x)\frac{\sqrt{\sigma_1^2+\sigma_2^2}}{\sqrt{2\pi}\sigma_1\sigma_2}\int_{\mathbb{R}}\exp\left(-\frac{\left(y\sqrt{\sigma_1^2+\sigma_2^2}-(x-\mu_1-\mu_2)\frac{\sigma_1^2}{\sqrt{\sigma_1^2+\sigma_2^2}}\right)^2}{2\sigma_1^2\sigma_2^2}\right)dy.
\end{aligned}$$

next use the linear transformation $y \mapsto \frac{y}{\sqrt{\sigma_1^2+\sigma_2^2}}$:

$$\begin{aligned}
&f_Z(x)\\
=&f_{\mu_1+\mu_2,\sigma_1^2+\sigma_2^2}(x)\frac{1}{\sqrt{2\pi}\sigma_1\sigma_2}\int_{\mathbb{R}}\exp\left(-\frac{\left(y-(x-\mu_1-\mu_2)\frac{\sigma_1^2}{\sqrt{\sigma_1^2+\sigma_2^2}}\right)^2}{2\sigma_1^2\sigma_2^2}\right)dy\\
=&f_{\mu_1+\mu_2,\sigma_1^2+\sigma_2^2}(x)\underbrace{\int_{\mathbb{R}}f_{(x-\mu_1-\mu_2)\frac{\sigma_1^2}{\sqrt{\sigma_1^2+\sigma_2^2}},\sigma_1^2\sigma_2^2}(y)dy}_{=1}=f_{\mu_1+\mu_2,\sigma_1^2+\sigma_2^2}(x).
\end{aligned}$$

This proves that $Z \sim \mathcal{N}(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$.

Finally, observe that we can directly compute $\mathbb{E}(Z)$ and $\mathrm{Var}(Z)$ by linearity and independence respectively,

$$\begin{aligned}
\mathbb{E}(Z) &= \mathbb{E}(X+Y) = \mathbb{E}(X)+\mathbb{E}(Y) = \mu_1+\mu_2,\\
\mathrm{Var}(Z) &= \mathrm{Var}(X+Y) = \mathrm{Var}(X)+\mathrm{Var}(Y) = \sigma_1^2+\sigma_2^2.
\end{aligned}$$

However, note that computation of mean (1st moment) and variance (2nd moment) is **not** enough to determine the distribution. In general, if one wants to identify uniquely the distribution of a bounded random variable, one needs to know all its moments.

Exercise 3. Let X be the number of network breakdowns that occur randomly and independently of each other on an average rate of 3 per month.

a Which model would you use to describe the phenomenon? Find the mean and variance of X.

Solution: The Poisson distribution is a suitable model to describe phenomenon:

$$X \sim \mathrm{Pois}(\lambda)$$

with parameter $\lambda = 3$. The PMF is given by

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

And we know that for the Poisson distribution $\mathbb{E}(X) = \mathrm{Var}(X) = \lambda$.

b What is the probability that there will be just 1 network breakdown in a month?

Solution:

$$\mathbb{P}(X = 1) = 3e^{-3} \approx 0.1494 = 14.94\%.$$

c What is the probability that there will be at least 6 network breakdowns in a month? Use `R` for this computation.

Solution:

$$\begin{aligned}
\mathbb{P}(X &\geq 6) = 1 - \mathbb{P}(X < 6) = 1 - \mathrm{CDF}(5) \\
&= 1 - \Big(\mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \cdots + \mathbb{P}(X = 5)\Big) \\
&= 1 - (1 + 3 + 3^2/2 + 3^3/6 + 3^4/24 + 3^5/120)e^{-3} \\
&\approx 0.083918 = 8.3918\%.
\end{aligned}$$

and using `R`:

```
(1- ppois(5, lambda=3))*100
[1] 8.391794
```

d In part a, you have found the mean and variance of X. Using only this information, apply *Chebyshev's inequality* to obtain a bound for $\mathbb{P}(X \geq 6)$ and compare the result with what you have found in part c.

Solution: By applying Chebyshev's inequality, we first show that for any random variable with $X \sim \mathrm{Pois}(\lambda)$ we have

$$\mathbb{P}(X \geq 2\lambda) \leq 1/\lambda.$$

Observe that

$$\begin{aligned}\mathbb{P}(X \geq 2\lambda) &\leq \mathbb{P}(X \geq 2\lambda) + \mathbb{P}(X \leq 0)\\ &= \mathbb{P}\big(X \geq 2\lambda \quad \cup \quad X \leq 0\big)\\ &= \mathbb{P}\big(X - \lambda \geq \lambda \quad \cup \quad X - \lambda \leq -\lambda\big)\\ &= \mathbb{P}(|X - \lambda| \geq \lambda).\end{aligned}$$

We know that for the Poisson distribution $\mathbb{E}(X) = \mathrm{Var}(X) = \lambda$. By Chebyshev's inequality

$$\mathbb{P}(|X - \lambda| \geq \lambda) = \mathbb{P}(|X - \mathbb{E}(X)| \geq \lambda) \leq \frac{\mathrm{Var}(X)}{\lambda^2} = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}.$$

Finally, in our example we have $\lambda = 3$. Therefore,

$$\mathbb{P}(X \geq 6) = \mathbb{P}(X \geq 2\lambda) \leq \frac{1}{\lambda} = \frac{1}{3} \approx 33.33\%$$

which is true for the exact computation in part c, $8.3918\% < 33.33\%$. As you see, the estimate above using Chebyshev's inequality is a very crude estimate.

Exercise 4. The yearly number of car accidents (denoted by X) in a city can be modeled by a Poisson distribution. In a given accident, the probability of a casualty is p. In this exercise, we want to find the distribution of the number of car accidents with casualties (denoted by Y). Let us consider $X \sim Pois(\lambda)$ and $Y|X \sim Binom(X; p)$ conditional upon X.

a Find the joint distribution of X and Y.

Solution: X and Y are discrete random variables. Let $n, m \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$. We have

$$\begin{aligned}\mathbb{P}(Y = m, X = n) &= \mathbb{P}(Y = m | X = n)\, \mathbb{P}(X = n)\\ &= \binom{n}{m} p^m (1-p)^{n-m} \times \frac{\lambda^n}{n!} \exp(-\lambda).\end{aligned}$$

If $m > n$, we use the convention that $\binom{n}{m} \equiv 0$.

b Prove that the marginal distribution of Y is given by $Y \sim Pois(p\lambda)$. (That is, the number of car accidents with casualties is again Poisson but with a smaller parameter.)

Solution: To find the marginal distribution of Y, we compute

$$
\begin{aligned}
\mathbb{P}(Y=m) &= \sum_{n=0}^{\infty} \mathbb{P}(Y=m, X=n) \\
&= \exp(-\lambda) \sum_{n=0}^{\infty} \binom{n}{m} p^m (1-p)^{n-m} \frac{\lambda^n}{n!} \\
&= \exp(-\lambda) \sum_{n=m}^{\infty} \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} \frac{\lambda^n}{n!} \\
\text{set } k := n-m \rightarrow \qquad &= \exp(-\lambda) \sum_{k=0}^{\infty} \frac{(m+k)!}{m!k!} p^m (1-p)^k \frac{\lambda^{m+k}}{(m+k)!} \\
&= \exp(-\lambda) \frac{p^m \lambda^m}{m!} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} (1-p)^k \\
&= \exp(-\lambda) \frac{p^m \lambda^m}{m!} \sum_{k=0}^{\infty} \frac{\big(\lambda(1-p)\big)^k}{k!} \\
&= \exp(-\lambda) \frac{p^m \lambda^m}{m!} \exp\big(\lambda(1-p)\big) = \frac{(\lambda p)^m}{m!} \exp(-\lambda p).
\end{aligned}
$$

So we conclude that $Y \sim Pois(\lambda p)$.

c Let $X' \sim Pois(\mu)$ be the yearly number of bicycle accidents, and assume that it is independent of X. Find the distribution of the total number of accidents $X + X'$. *Hint*: use formula (1).

Solution: X and X' are independent and discrete random variables. Thus, we can apply formula (1) to find PMF of the random variable $Z = X + X'$

$$
\begin{aligned}
\mathbb{P}(Z=k) &= \sum_{n \in \mathbb{N}_0} \mathbb{P}(X=n)\mathbb{P}(X'=k-n) \\
&= \sum_{n \in \mathbb{N}_0} \frac{\lambda^n}{n!} \exp(-\lambda) \times \frac{\mu^{k-n}}{(n-k)!} \exp(-\mu) \\
&= \exp(-(\lambda+\mu)) \frac{1}{k!} \sum_{n \in \mathbb{N}_0} \frac{k!}{n!(n-k)!} \lambda^n \mu^{k-n} \\
&= \exp(-(\lambda+\mu)) \frac{1}{k!} (\lambda+\mu)^k.
\end{aligned}
$$

$$
\implies X + X' \sim Pois(\lambda+\mu).
$$

d What is the distribution of the number of bicycle accidents if we know that the total number accidents in a year is k?

Solution: We know $X \sim Pois(\lambda)$ and $X' \sim Pois(\mu)$ are independent.

$$\begin{aligned}
\mathbb{P}(X' =& m|X + X' = k) \\
&= \frac{\mathbb{P}(X' = m, X + X' = k)}{\mathbb{P}(X + X' = k)} \\
&= \frac{\mathbb{P}(X' = m, X = k - m)}{\mathbb{P}(X + X' = k)} \\
&= \frac{\mathbb{P}(X' = m)\mathbb{P}(X = k - m)}{\mathbb{P}(X + X' = k)} \\
&= \frac{\mu^m}{m!}\exp(-\mu) \cdot \frac{\lambda^{k-m}}{(k-m)!}\exp(-\lambda) \Big/ \left(\frac{(\lambda+\mu)^k}{k!}\exp(-(\lambda+\mu))\right) \\
&= \frac{k!}{m!(k-m)!}\frac{\mu^m\lambda^{k-m}}{(\lambda+\mu)^k} \\
&= \binom{k}{m} q^m (1-q)^{k-m}, \qquad q := \frac{\mu}{\mu+\lambda}.
\end{aligned}$$

$$\implies X'|(X + X') \sim Binom(X + X', \frac{\mu}{\mu+\lambda}).$$

We conclude that the conditional distribution of a Poisson r.v. on its sum with another independent Poisson r.v. is binomial.

Exercise 5. The exponential distribution $Exp(\lambda)$ with rate parameter $\lambda > 0$ is typically used to model the waiting time $X \geq 0$ until the occurrence of a certain event. Then $\mathbb{E}(X) = 1/\lambda$ is the average time until the occurrence of the event of interest (measured in some given unit of time).

A crucial property of the exponential distribution is that it is "*memoryless*": No matter how long you have been waiting already, the probability of waiting for an additional amount of time $s > 0$ only depends on s, and not on your past waiting time $t > 0$. This can be written as

$$\mathbb{P}(X > t + s|X > t) = \mathbb{P}(X > s). \tag{3}$$

Prove identity (3) using the CDF of $X \sim Exp(\lambda)$.

Solution: Let us first recall that

$$F_X(t) := \mathbb{P}(X \leq t) = 1 - e^{-\lambda t}, \quad t \geq 0.$$

We use definition of conditional probability:

$$\begin{aligned}\mathbb{P}(X > t + s | X > t) &= \frac{\mathbb{P}(X > t + s \cap X > t)}{\mathbb{P}(X > t)} \\ &= \frac{\mathbb{P}(X > t + s)}{\mathbb{P}(X > t)} \\ &= \frac{1 - \mathbb{P}(X \leq t + s)}{1 - \mathbb{P}(X \leq t)} \\ &= \frac{e^{-\lambda(t+s)}}{e^{-\lambda(t)}} \\ &= e^{-\lambda(s)} \\ &= \mathbb{P}(X > s).\end{aligned}$$

Exercise 6. The **Pearson correlation coefficient** (cf. Def. 6 in Ch. 1.5) of two random variables X and Y (with $\mathbb{E}(X^2)$, $\mathbb{E}(Y^2) < \infty$) is defined to be 0 if $\mathrm{Var}(X) = 0$ or $\mathrm{Var}(Y) = 0$, and otherwise

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}}.$$

Prove that the Pearson coefficient always satisfies

$$-1 \leq \rho(X, Y) \leq 1,$$

with the equality if and only if there is a *linear relationship* between X and Y. Namely,

$$|\rho(X, Y)| = 1 \iff Y = cX + d,$$

where

$$c = \begin{cases} \sqrt{\frac{\mathrm{Var}(Y)}{\mathrm{Var}(X)}}, & \rho(X, Y) = 1, \\ -\sqrt{\frac{\mathrm{Var}(Y)}{\mathrm{Var}(X)}}, & \rho(X, Y) = -1, \end{cases} \quad , \quad d = \mathbb{E}(Y) - c\mathbb{E}(X).$$

Hint: use the **Cauchy–Schwarz inequality** (cf. Corollary (2) in Ch. 1.4)

$$|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2)} \cdot \sqrt{\mathbb{E}(Y^2)}$$

for any $X, Y : \Omega \to \mathbb{R}$ (with $\mathbb{E}(X^2)$, $\mathbb{E}(Y^2) < \infty$), whereas the equality holds if and only if $X = aY$ for some constant $a \in \mathbb{R}$.

Solution: By definition of covariance (cf. Def. 5 in Ch. 1.5) we have

$$\operatorname{Cov}(X,Y) = \mathbb{E}\left[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right].$$

Define $\hat{X} := X - \mathbb{E}(X)$ and $\hat{Y} := Y - \mathbb{E}(Y)$ and apply Cauchy–Schwarz inequality for $\hat{X}$ and $\hat{Y}$:

$$|\operatorname{Cov}(X,Y)| = |\mathbb{E}(\hat{X}\hat{Y})| \leq \sqrt{\mathbb{E}(\hat{X}^2)} \cdot \sqrt{\mathbb{E}(\hat{Y}^2)} = \sqrt{\mathrm{Var(X)}} \cdot \sqrt{\mathrm{Var(Y)}}. \quad (4)$$

This proves that $-1 \leq \rho(X,Y) \leq 1$.

Next, note that according to the hint above, we will have the equality in (4) (i.e. $|\rho(X,Y)| = 1$) if and only if $\hat{Y} = c\hat{X}$ for some constant $c \in \mathbb{R}$. This implies

$$Y = cX + \big(- c\mathbb{E}(X) + \mathbb{E}(Y)\big).$$

and in this case

$$\operatorname{Var}(Y) = \operatorname{Var}(\hat{Y}) = \operatorname{Var}(c\hat{X}) = c^2 \operatorname{Var}(\hat{X}) = c^2 \operatorname{Var}(X) \quad \Longrightarrow c = \pm\sqrt{\frac{\operatorname{Var}(Y)}{\operatorname{Var}(X)}}$$

$$\operatorname{Cov}(X,Y) = \mathbb{E}(\hat{X}\hat{Y}) = c\mathbb{E}(\hat{X}^2) = c\operatorname{Var}(X) \implies \rho(X,Y) = \pm 1.$$