1. Uvažujme posloupnost n znaků, ve které se každý znak zkreslí s pravděpodobností $p \in (0,1)$, a to nezávisle na ostatních znacích. Nechť X značí počet takto zkreslených znaků. Dále označme jako Y_i identifikátor toho, zda se i-tý znak zkreslí.

   (a) Připomeňte si, jaké rozdělení má X a jakou má střední hodnotu a rozptyl.

   (b) Vyjádřete X pomocí veličin $Y_1, \dots, Y_n$ a pomocí tohoto vyjádření spočtěte $\mathsf{E}X$ a $\mathsf{Var}\, X$.

Y_i .. zkreslí se s p

$$EY_i = 0 \cdot (1-p) + 1p = p$$

$$X = \sum_i Y_i$$

$$EX = \sum_i EY_i = np \qquad X \sim Bi(n,p)$$

$$\mathrm{var}\, Y_i = p - p^2 = p(1-p)$$

$$\mathsf{Var}\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 \mathsf{Var}\, X_i + \sum\sum_{i \neq j} a_i a_j \mathsf{Cov}\,(X_i, X_j)$$

1

0 protože nezávislé

$$\mathrm{var}\, X = \sum \mathrm{var}\, Y_i = np(1-p)$$

2. V rybníce plave 100 rybiček, z nichž je a zlatých. Náhodně vylovíme 5 rybiček a náhodná veličina X udává, kolik zlatých rybiček jsme vylovili.

   (a) Vyjádřete analogicky jako v 1. příkladě veličinu X jako součet 0-1 náhodných veličin Y_i. Určete rozdělení Y_i.

   (b) Spočtěte očekávanou hodnotu X.

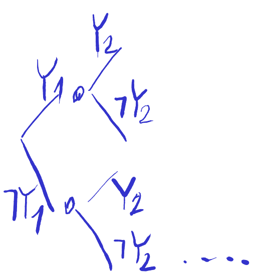
   (c) Jsou veličiny $Y_1, \dots, Y_n$ nezávislé? Pokud nejsou, spočtěte jejich kovarianci.

   (d) Spočtěte rozptyl X.

$$Y_i = \begin{cases} 1 & \text{pokud jsme v } i\text{-tém tahu vytáhli zlatou} \\ 0 & \text{jinak} \end{cases}$$

$$P[Y_1 = 1] = \frac{a}{100} \qquad P[Y_1 = 0] = \frac{100-a}{100}$$

$$P[Y_2 = 1] = P[Y_2 = 1 \mid Y_1 = 0] + P[Y_2 = 1 \mid Y_1 = 1]$$



$$= \frac{a}{99} \cdot \frac{(100-a)}{100} + \frac{a-1}{99} \cdot \frac{a}{100} =$$

$$= \frac{100a - a^2 + a^2 - a}{99 \cdot 100} = \frac{99a}{99 \cdot 100} = \frac{a}{100}$$

$$P[Y_i = 1] = \frac{a}{100} \qquad EY_i = \frac{a}{100}$$

$$EX = \sum_{i=1}^{5} EY_i = 5 \cdot \frac{a}{100} = \frac{a}{20}$$

c) NEJSOU NEZÁVISLÉ

$$P[Y_i = 1] = \frac{a}{100}$$

$$\operatorname{var} Y_i = \underbrace{p(1-p)}_{\text{alternativní rozdělení}} = \frac{a}{100} \cdot \frac{(100-a)}{100} = \frac{a(100-a)}{10000}$$

$$\operatorname{cov}(Y_i, Y_j) = EY_iY_j - EY_iEY_j$$

$$EY_iY_j = 0 \cdot P[Y_i=0, Y_j=0] + 0 \cdot P[Y_i=0, Y_j=1] + 0 \cdot P[Y_i=1, Y_j=0] + 1 \cdot P[Y_i=1, Y_j=1]$$

$$= 1 \cdot \frac{a}{100} \cdot \frac{a-1}{99}$$

$$\Big(P[Y_1=1, Y_2=1] = \frac{a}{100} \cdot \frac{a-1}{99}$$

$$P[Y_1=1, Y_3=1] = P[Y_1=1, Y_2=0, Y_3=1] + P[Y_1=1, Y_2=1, Y_3=1]$$

$$= \frac{a(100-a)(a-1)}{100 \cdot 99 \cdot 98} + \frac{a(a-1)(a-2)}{100 \cdot 99 \cdot 98} =$$

$$= \frac{(a-1)a}{100 \cdot 99} \underbrace{\left(\frac{100-a+a-2}{98}\right)}_{1} = \frac{a(a-1)}{100 \cdot 99}$$

$$\operatorname{cov}(Y_i, Y_j) = \frac{a(a-1)}{100 \cdot 99} - \frac{a}{100} \cdot \frac{a}{100} =$$

$$= \frac{a}{100}\left(\frac{a-1}{99} - \frac{a}{100}\right)$$

$$= \frac{a}{100}\left(\frac{100(a-1) - 99a}{100 \cdot 99}\right) = \frac{a}{100}\left(\frac{a-100}{100 \cdot 99}\right) < 0$$

$$\operatorname{corr}(Y_i, Y_j) = \frac{\frac{a(a-100)}{100 \cdot 100 \cdot 99}}{\frac{a(100-a)}{100^2}} = -\frac{1}{99}$$

$$\operatorname{Var}\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 \operatorname{Var} X_i + \sum\sum_{i \neq j} a_i a_j \operatorname{Cov}(X_i, X_j)$$

$$\operatorname{var}(X) = \sum_{i=1}^{5} \operatorname{var} Y_i + \sum\sum_{i \neq j} \operatorname{cov}(Y_i, Y_j) = 5 \cdot \frac{a(100-a)}{10000} + \binom{5}{2} \frac{a}{100}\left(\frac{a-100}{100 \cdot 99}\right) = \frac{a(a-100)}{100^2}\left(5 - \frac{10}{99}\right)$$

3. Hmotnosti jednotlivých zásilek na poště jsou nezávislé náhodné veličiny, které mají všechny stejné rozdělení se střední hodnotou μ a rozptylem σ^2. Rozhodli jsme se změřit hmotnosti n zásilek a jsou to náhodné veličiny $X_1, \ldots, X_n$.

   (a) Označme jako $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ průměrnou hmotnost zásilky. Spočtěte střední hodnotu a rozptyl této náhodné veličiny. Jak se bude rozptyl měnit, budeme-li zvyšovat n?

   Předpokládejme dále, že X_i mají rovnoměrné rozdělení na intervalu $[a, b]$, kde $0 < a < b$ jsou reálná čísla. Označme jako $M_n = \max_{i=1,\ldots,n} X_i$ maximální hmotnost a $L_n = \min_{i=1,\ldots,n} X_i$ minimální hmotnost.

   (b) Vyjádřete distribuční funkci a hustotu náhodné veličiny M_n.
   (c) Spočtěte očekávanou hodnotu M_n. Jak se bude tato hodnota chovat, budeme-li navyšovat počet pozorování n?
   (d) Spočtěte dále rozptyl M_n.
   (e) Vyjádřete rozdělení a spočtěte očekávanou hodnotu L_n.

$X_i \sim U[a,b] \qquad M_n = \max_i X_i \qquad L_n = \min_i X_i$

$$F_{M_n}(m) = P[M_n \le m] = P[X_i \le m \ \forall i] = P\left(\bigcap_{i=1}^{n} [X_i \le m]\right) = \left(P[X_i \le m]\right)^n = \left(F_X(m)\right)^n$$

$$f_{M_n}(m) = n\left(F_X(m)\right)^{n-1} \cdot f_X(m)$$

$$f_X(m) = \frac{1}{b-a}$$

$$F_X(m) = \frac{m-a}{b-a} \qquad m \in (a-b)$$

$$f_{M_n}(m) = n\left(\frac{m-a}{b-a}\right)^{n-1} \cdot \frac{1}{b-a} = \frac{1}{(b-a)^n} \cdot n \cdot (m-a)^{n-1}$$

$$E(M_n) = \int_a^b m \frac{1}{(b-a)^n} \cdot n \cdot (m-a)^{n-1}\, dm = \frac{1}{(b-a)^n} \cdot n \cdot \int_a^b m\,(m-a)^{n-1}\, dm = \left| \begin{array}{l} v' = (m-a)^{n-1} \\ v = \frac{(m-a)^n}{n} \\ u = m \\ u' = 1 \end{array} \right.$$

$$= \frac{n}{(b-a)^n}\left(\left[m \cdot \frac{(m-a)^n}{n}\right]_a^b - \int_a^b \frac{(m-a)^n}{n}\, dm\right) =$$

$$= \frac{n}{(b-a)^n}\left(\frac{(b-a)^n}{n} \cdot b - \left[\frac{(m-a)^{n+1}}{n \cdot (n+1)}\right]_a^b\right) = b - \frac{b-a}{n+1}$$

$$L_n : \quad F_{L_n}(m) = P[\exists i : X_i \le m]$$

$$= 1 - P[X_i > m \ \forall i] = 1 - \underbrace{P[X_1 > m, X_2 > m, \ldots]}_{\text{nezávislé}} =$$

$$= 1 - \prod_{i=1}^{n} P[X_i > m] = 1 - \left(1 - P[X_1 \le m]\right)^n =$$

$$= 1 - \left(1 - F_X(m)\right)^n$$

$$F_{L_n} = 1 - \left(\frac{b-m}{b-a}\right)^n \qquad E(L_n) = a + \frac{b-a}{n+1} \ ? \qquad f_{L_n} = \frac{1}{(b-a)^n}\, n\, (b-m)^{n-1}$$

5. Předpokládejme, že počet automobilových nehod v daný den na dálnici D1 je náhodná veličina X_i, která se řídí Poissonovým rozdělením s parametrem $\lambda > 0$, tj. $P(X_i = k) = e^{-\lambda}\lambda^k/k!$ for $k = 0, 1, \ldots$. Dále předpokládejme, že počty nehod v jednotlivých dnech jsou nezávislé náhodné veličiny.

   (a) Spočtěte očekávaný počet nehod v daný den.

   (b) Vyjádřete rozdělení a střední hodnotu počtu nehod za dva po sobě jdoucí dny.

a) $$EX_i = \sum_{k=1}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \underbrace{\sum_{j=0}^{\infty} \frac{\lambda^j}{j!}}_{\sum_{j=0}^{\infty} \frac{x^j}{j!} = e^x} = \lambda e^{-\lambda} e^{\lambda} = \underline{\underline{\lambda}}$$

b) $Z = X_1 + X_2$

$$P(Z=k) = P(X_1 + X_2 = k)$$

$$= \sum_{i=0}^{k} P(X_1 = i, X_2 = k-i) \overset{\text{nezávislost}}{=} \sum_{i=0}^{k} P(X_1 = i) \cdot P(X_2 = k-i)$$

$$= \sum_{i=0}^{k} e^{-\lambda} \frac{\lambda^i}{i!} \cdot e^{-\lambda} \frac{\lambda^{k-i}}{(k-i)!} = e^{-2\lambda} \sum_{i=0}^{k} \frac{\lambda^{i+(k-i)}}{i!\,(k-i)!} =$$

$$= e^{-2\lambda} \frac{1}{k!} \sum_{i=0}^{k} \frac{\lambda^k\, k!}{i!\,(k-i)!} = e^{-2\lambda} \frac{1}{k!} \underbrace{\sum_{i=0}^{k} \binom{k}{i} \lambda^{k-i} \cdot \lambda^i}_{\text{binomická věta}} =$$

$$= e^{-2\lambda} \frac{1}{k!} (\lambda + \lambda)^k$$

$$= e^{-2\lambda} \frac{(2\lambda)^k}{k!}$$

$Z \sim Exp(2\lambda)$

$EZ = 2\lambda$

6. Uvažujme nezávislé náhodné veličiny X, Y. Nalezněte rozdělení součtu $Z = X + Y$, pokud X i Y mají obě

(a) binomické rozdělení $\mathsf{Bi}(n, p)$ s parametry $n \geq 1$ a $p \in (0, 1)$,
(b) rovnoměrné rozdělení na intervalu $[0, 1]$,
(c) exponenciální rozdělení s hustotou $f(x) = \lambda e^{-\lambda x} I(x \geq 0)$ s parametrem $\lambda > 0$.

a) $X \sim Bi(n,p)$: $P[X=k] = \binom{n}{k} p^k (1-p)^{n-k}$

$Z = X + Y$

$$P(Z=k) = P(X+Y=k) = \sum_{i=0}^{k} P(X=i, Y=k-i) =$$

nezávislost

$$= \sum_{i=0}^{k} P(X=i)\, P(Y=k-i) =$$

$$= \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \cdot \binom{n}{k-i} p^{k-i} (1-p)^{n-(k-i)}$$

$$= \sum_{i=0}^{k} \binom{n}{i}\binom{n}{k-i} p^k (1-p)^{2n-k}$$

In combinatorics, **Vandermonde's identity**
coefficients:

$$\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k}$$

$$= \binom{2n}{k} p^k (1-p)^{2n-k}$$

$Z \sim Bi(2n, p)$

6. Uvažujme nezávislé náhodné veličiny X, Y. Nalezněte rozdělení součtu $Z = X + Y$, pokud X i Y mají obě

(a) binomické rozdělení $\mathsf{Bi}(n, p)$ s parametry $n \geq 1$ a $p \in (0, 1)$,
(b) rovnoměrné rozdělení na intervalu $[0, 1]$,
(c) exponenciální rozdělení s hustotou $f(x) = \lambda e^{-\lambda x} I(x \geq 0)$ s parametrem $\lambda > 0$.

b) $X \sim U_{[0,1]}$ $\quad Y \sim U_{[0,1]}$ $\quad Z = X + Y$

$$f_X(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & \text{jinak} \end{cases}$$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx$$

čísla z můžou být z $[0,2]$, jinde 0

I) $0 \leq z \leq 1$: $f_Y(z-x) > 0 \Rightarrow 1 > z - x > 0 \Rightarrow z > x \Rightarrow x \in (0, z)$

nikdy

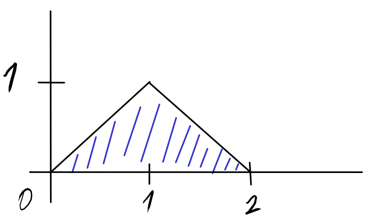
$$\int_0^z 1\, dx = [x]_0^z = z$$

II) $1 \leq z \leq 2$ $\quad f_Y(z-x) > 0 \Rightarrow 1 > z - x > 0 \Rightarrow x > z - 1$

nikdy

$$\int_{z-1}^{1} 1\, dx = [x]_{z-1}^{1} = 1 - (z-1) = 2 - z$$

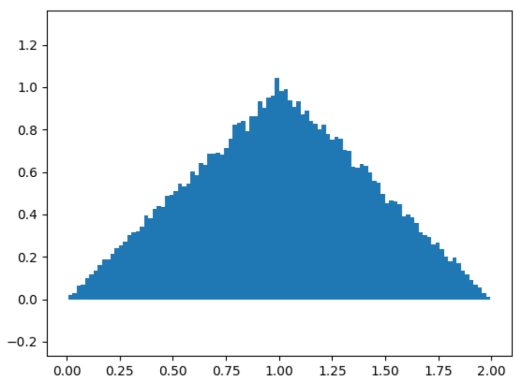
$$f_Z(z) = \begin{cases} z & \text{pro } z \in (0,1) \\ 2 - z & \text{pro } z \in (1,2) \\ 0 & \text{jinak} \end{cases}$$




```python
def pokus():

    size_gen = 100000
    a = np.random.uniform(low=0, high=1, size=size_gen)
    b = np.random.uniform(low=0, high=1, size=size_gen)
    c = a + b

    plt.hist(c, bins=100, density=True)
    plt.axis("equal")
    plt.show()
```



c) viz 5.