Assignment 8.

N1. An average of 1 typo per page in a 600-page book means there's 600 typos in the book.

Let T_{15} be the amount of typos on page 13, $T_{15} \sim Bin(600, \frac{1}{600})$ Let T_{15} be the amount of typos on page 13, $T_{15} \sim Bin(600, \frac{1}{600})$

Let
$$I_{15}$$
 be the amount of types on page 13, I_{15} be the amount of types on page 13, I_{15} be the amount of types on page 13, I_{15} be the amount of types on page 13, I_{15} be the amount of types on page 13, I_{15} be the amount of types of page 13, I_{15} be the amount of types of page 13, I_{15} be the amount of types of page 13, I_{15} be the amount of types of page 13, I_{15} be the amount of types of page 13, I_{15} be the amount of types 13, I_{15} be the amount of types 14, I_{15} be the amount of types 15, I_{15} be the a

```
Nz. Let p=0,003 be the probability that a scane doesn't have any raisins.
 Let S be the amount of scones without raisins within a batch of 1000 scones.
S~Bin (1000, 0,003), \ = 1000.0,003 = 3
Therefore, P(S=k) \approx \frac{\lambda}{k!}e^{-\lambda} = \frac{3}{k!}e^{-\lambda}
 a) P(S=0) \approx \frac{3}{0!e^3} = \frac{1}{e^3} = 0.05
B) P(S=3) \approx \frac{3^3}{3!e^5} = \frac{27}{6e^3} = 0,224
C) P(S \ge 3) = 1 - P(S < 3) = 1 - (P(S = 0) + P(S = 1) + P(S = 2)) \approx 1 - (0.05 + \frac{3}{e^{T}} + \frac{9}{2e^{1}}) = 0.576
N3. Let R be the amount of raisins in a scone.
Assuming there is x raisins among 1000 scones, we have R~Bin(x, 1000)
\lambda = x \cdot \frac{1}{1000} = ER - the average amount of raisins in a scone.
 P(R=0) \approx \frac{\lambda^{\circ}}{o!e^{\lambda}} = \frac{1}{e^{\lambda}}, we want \lambda: \frac{1}{e^{\lambda}} \leq 0.01 \Rightarrow \lambda \geqslant -\ln(0.01) = 4.6
Therefore, we need 4,6 raisins per scone on average.
N4. Let N=5000 be the amount of electors, P=0,7, q=0,3
Let 3 be the amount of electors who voted for candidate A, then N-3 people voted for B.
3~ Bin (N, p), 3 = \(\sum_{i=1}^{2}\); , where \(\frac{2}{2}\); ~ Bin(1,p) - indicator that the i-th elector voted for A.
Since 3: are independent and identically distributed, one can apply the Central Limit Theorem:
P(3-E1 < y) = P(3-NP < y) = P(3 < x) = P(3 <
a) P(3-(N-3)=1900)=P(3=\frac{1900+N}{2})=P(3=3450)\approx P(3\leq 3450,5)-P(3\leq 3449,5)\approx
                                                                           \approx \Phi\left(\frac{3450,5-Np}{\sqrt{Npq'}}\right) - \Phi\left(\frac{3449,5-Np}{\sqrt{Npq'}}\right) = \Phi(-1,527) - \Phi(-1,558) = 0,0037
 B) P(3-(N-3) ≥ 1900) ≈ P(µ≥-1,54) = 1- P(-1,54) = 1-0,063 = 0,937
N5. Let 3 be the amount of people that went in through the first entrance.
  Case (a): \frac{9}{4}~ 2 \sin(500, \frac{1}{2}), E_a^9 = 2.500.\frac{1}{2} = 500, Var \frac{9}{4} = 2^2.500.\frac{1}{2}.\frac{1}{2} = 500
 Case (b): 30~ Bin (1000, 1), E 30= 1000. 1= 500, Var 30= 1000. 1.1= 250
   3a = \( \sum_{i=1}^{500} \gamma_{ai}\), where \( \frac{3}{2} \) = i.i.d. \( \frac{5}{2} \) = for \( \frac{9}{a} \) and \( \frac{9}{8} \) one can apply the Central Limit Theorem
  36 = \( \sum_{i=1}^{1000} \frac{9}{6}; \) where \( \frac{9}{6}; \sup \text{Bin}(1, \frac{1}{2}) - i.i.d. \)
   P(3-E1a < y) = P(3-500 < y) = P(4) => P(3a < y) = P(3a
  Assuming that ka and ke are the amounts of doakroom places in cases (a) and (8),
   one needs to find ka, kg:
                                                                                          Note: since \frac{90}{2} and se are binomially distributed, and p=\frac{1}{2}=1-p,
       P(sa & ka) = 0,99
                                                                                                             P(se = ke) = 0,99 implies P(1000 - se = ke) = 0,99, since
      P(3e = ke) = 0,99
                                                                                                              P(98 = x) = P(98 > 1000 -x) Y x E[0; 1000]
```

differs from the answer because 2,575 was not rounded to 2,57

$$P(\S_a \le k_a) \approx P(\frac{k_a - 500}{\sqrt{500^7}}) = 0.99$$

$$P(\S_a \le k_a) \approx P(\frac{k_a - 500}{\sqrt{500^7}}) = 0.99$$

$$\frac{k_a - 500}{\sqrt{500}} = 2.575$$

$$\frac{k_a - 500}{\sqrt{500}} =$$

N6. Since a coin is fair, the probability of getting tails is p=0.5 Let g_k be the percentage of tails among k throws (expressed as a vatio) Then $g_k \sim \frac{1}{k} \sin(k, \frac{1}{2})$, $E g_k = \frac{1}{k} \cdot k \cdot \frac{1}{2} = \frac{1}{2}$, $Var g_k = (\frac{1}{k})^2 \cdot k \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4k}$

One needs to find min k: P(p-0,01 = 3, < p+0,01) = 0,95

P(\$ = [0,49; 0,51]) = P(0,51-1) - P(0,49-1) = P(0,025k) - P(-0,025k) = 0,95

 $\Phi(x) - \Phi(-x) = (0.5 + \Phi_0(x)) - (0.5 + \Phi_0(-x)) = \Phi_0(x) - \Phi_0(-x) = 2\Phi_0(x)$, since $\Phi_0(x) = 0$

Thus, 240(0,02 Tk) = 0,95 => 4(0,02 Jk) = 0,5+ 0,95 = 0,975

According to the standard normal table,

P(0,02√k) = 0,975 ⇔ 0,02√k = 1,96 ⇒ k = 9604

Therefore, the minimum amount of flips is 9604.

N7. Since a die is fair, the probability of getting a "4" is &.

Let g_k be the percentage of "4"s among k rolls (expressed as a ratio) Then $g_k \sim \frac{1}{k} \text{Bin}(k, \frac{1}{6})$, $Eg_k = \frac{1}{k} \cdot k \cdot \frac{1}{6} = \frac{1}{6}$, $Var g_k = (\frac{1}{k})^2 \cdot k \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{5}{36k}$

One needs to find min ka: $P(\frac{1}{6}-0.01 \le 3k_a \le \frac{1}{6}+0.01) = 0.95$

P(8ke[1-901; 1+9,01]) = P(9,01) - P(-9,01) = 0,95

As shown in the previous problem, $P(3_{ka} \in [\frac{1}{6}-0.01; \frac{1}{6}+0.01]) = 2 P_0(\frac{0.06}{\sqrt{5}} \sqrt{k_a})$

According to the standard normal table, $P(\frac{0.06}{55} \text{ Jka}) = 0.5 + \frac{0.95}{2}$ for $\frac{0.06}{55} \text{ Jka} = 1.96$ Thus, $k_a = (\frac{1.965}{0.06})^2 = 5335, 5$ the minimum amount of rolls is [5335, 5] = 5336

One also needs to find min k_8 : $P(\frac{1}{6}-0.01 \le 3k_8 \le \frac{1}{6} + 0.01) > 0.99$ $2\Phi_0(\frac{0.06}{\sqrt{5}}) > 0.99$ for $\frac{0.06}{\sqrt{5}}$ $\sqrt{k_8} > 2.575$ $(\Phi : 5.7)$

Thus, $k_B > (\frac{2,575}{0,06})^2 = 9209, 2$ so the minimum amount of rolls is $\lceil 9209, 2 \rceil = 9210$

N8. The time to cycle 3 miles on a wet day is $\frac{3}{10} = 0.3$ hours $\frac{3}{20} = 0.15$ hours

Let T; be the time to cycle on the i-th day of the course.

 $T_i \sim \begin{pmatrix} 0,3 & 0,15 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ $\forall i \implies ET_i = \frac{0,45}{2} = 0,225 \text{ hours}$

Let T be the total cycling time over the whole course.

 $T = \sum_{i=1}^{24} T_i \Rightarrow ET = \sum_{i=1}^{24} ET_i = 24 \cdot \frac{0.45}{2} = 12.0,45 = 5,4 \text{ hours}$

$$\frac{N9}{N9}$$
. Let $X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$, $Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_m \end{bmatrix}$ be independent multivariate normal

Let
$$\overline{\mu}_{x} = \begin{bmatrix} \mu_{1} \\ \mu_{1} \end{bmatrix}$$
 and $\Sigma_{x} = \begin{bmatrix} \sigma_{1}^{2} & 0 \\ 0 & \sigma_{1}^{2} \end{bmatrix}$ be the mean vector and covariance matrix of \overline{X} , $\overline{\mu}_{y} = \begin{bmatrix} \mu_{1} \\ \mu_{2} \end{bmatrix}$ and $\Sigma_{y} = \begin{bmatrix} \sigma_{1}^{2} & 0 \\ 0 & \sigma_{1}^{2} \end{bmatrix}$ be the mean vector and covariance matrix of \overline{Y}

$$\sum_{x} [i,j] = \sum_{y} [i,j] = 0$$
 for $i \neq j$ since X_i and Y_i are all mutually independent.

$$f_{x}(x_{n},x_{2},...,x_{n}) = \frac{1}{(\sqrt{2\pi})^{n}\sqrt{\det\Sigma_{x}}}e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{1}{\sigma_{i}^{2}}(x_{i}-\mu_{i})^{2}}, \quad f_{y}(y_{1},y_{2},...,y_{m}) = \frac{1}{(\sqrt{2\pi})^{m}\sqrt{\det\Sigma_{y}}}e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{1}{\sigma_{i}^{2}}(y_{i}-\mu_{i})^{2}}$$

Since X and Y are independent,
$$f_{x,y}(x_1,...,x_n,y_1,...,y_m) = f_{x}(x_1,...,x_n) \cdot f_{y}(y_1,...,y_m)$$

Then we can introduce $Z = \begin{bmatrix} X \\ Y \end{bmatrix}$, $\overline{M}_{z} = \begin{bmatrix} M_{1} \\ M_{2} \end{bmatrix}$, $\sum_{z} = \begin{bmatrix} \sigma_{1}^{z} & \sigma_{2}^{z} \\ \sigma_{1}^{z} & \sigma_{2}^{z} \end{bmatrix}$

Z is multivariate normal since
$$f_{z}(\bar{z}) = f_{x,y}(\bar{x},\bar{y}) = \frac{1}{(\sqrt{2\pi})^{n+m}\sqrt{det}\Sigma_{x}det}\sum_{y} e^{-\frac{1}{2}(\sum_{j=1}^{n}(x_{i}-y_{j})^{2}+\sum_{j=1}^{n}\frac{1}{\sigma_{y}^{2}}(y_{j}-y_{j})^{2})}$$

According to the properties of a multivariate normal distribution, any linear combination of Z with coefficients c.,..., cmm is normally distributed like:

Therefore,
$$\bar{X} - \bar{Y} = \sum_{i=1}^{n} \frac{1}{n} X_i - \sum_{j=1}^{n} \frac{1}{m} Y_j \sim N(n \cdot \frac{1}{n} \mu_1 - m \cdot \frac{1}{m} \mu_2; n(\frac{1}{n})^2 \cdot \sigma_1^2 + m \cdot (-\frac{1}{m})^2 \cdot \sigma_2^2)$$

$$\sim N(\mu_1 - \mu_2; \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m})$$

Let
$$U=X-Y$$
, then $f_u(t)=\frac{1}{\sqrt{2\pi(\frac{n}{n}+\frac{n}{m})}}e^{-\frac{1}{2}\cdot\frac{1}{\frac{n}{n}+\frac{n}{m}}(t-\mu_1+\mu_2)^2}$

N10. Let X, Y~N(0; 1) be independent RVs

Treating them as univariate normal, we have that Z=[Y] is bivariate normal,

According to the properties of a multivariate normal distribution, a vector composed of linear combinations of a multivariate normal vector is also multivariate normal:

$$W = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \times Z = \begin{bmatrix} x + y \\ x - y \end{bmatrix}, \quad W \sim N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}; \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix})$$

$$f_{w}(x, y) = \frac{1}{2\pi \cdot 2} e^{-\frac{x^{2} + y^{2}}{4}}$$

$$\begin{cases} x + y = \begin{bmatrix} 1 & 1 \end{bmatrix} \times Z, \quad X - Y = \begin{bmatrix} 1 & -1 \end{bmatrix} \times Z, \quad \text{both are normally distributed:} \end{cases}$$

$$f_{x \cdot y}(u) = \frac{1}{\sqrt{4\pi}} e^{-\frac{u^{2}}{4}} \qquad , \qquad f_{x \cdot y}(v) = \frac{1}{\sqrt{4\pi}} e^{-\frac{v^{2}}{4}}$$

$$f_{w}(x, y) = f_{x \cdot y}(x) \cdot f_{x \cdot y}(y) \implies X + Y \text{ and } X - Y \text{ are independent}$$

$$Let Z' = |Y| \cdot (-1)^{I_{x \cdot 0}} = \begin{cases} |Y|, \quad x > 0 \\ -|Y|, \quad x < 0 \end{cases}$$

$$P(Z' < t) = P(x > 0) \cdot P(|Y| < t) + P(x < 0) \cdot P(|Y| > -t) = 0, s(P(|Y| < t) + P(|Y| > -t))$$

$$P(Z' < t) = \begin{cases} 0, s(P(|Y| < t) + 1), \quad t > 0 \end{cases} \qquad (1)$$

$$Q_{1} \leq (P(|Y| > -t)), \qquad t < 0 \qquad (2)$$

$$(1) = 0.5 + 0.5P(-t < Y < t) = 0.5 + 0.5P(t) - 0.5P(-t) = 0.5 + 0.5 \cdot (0.5 + P_0(t)) - 0.5(0.5 + P_0(-t)) = 0.5 + 0.5 \cdot (0.5 + P_0(t)) - 0.5(0.5 + P_0(-t)) = 0.5 + 0.5 \cdot (0.5 + P_0(t)) - 0.5P(-t) = 0.5 \cdot (0.5 + P_0(t)) - 0.5P(-t) = 0.5 \cdot (0.5 + P_0(t)) = 0.5 \cdot (0.5 + P_0(t$$

To show that
$$\begin{bmatrix} x \\ z' \end{bmatrix}$$
 is not bivariate normal, lets consider the joint CDF:
$$P(X < u, -|Y| < v) \qquad , \qquad u < 0$$

$$P(X < u, Z' < v) = \begin{cases} P(x < u, -|Y| < v) + P(X < 0, -|Y| < v), & u > 0 \end{cases}$$

$$\begin{cases} P(x < u)P(|Y| > -v), & u < c, v > 0 \\ P(x < u)P(Y > -v) + P(x < u)P(Y < v), & u < c, v < 0 \\ P(x < c)P(Y > -v) + P(x < c)P(Y < v), & u > 0, v < 0 \\ P(x < c)P(|Y| > -v) + P(x < c)P(Y < v), & u > 0, v < 0 \\ P(x < c)P(|Y| > -v) + P(c < x < u)P(-v < Y < v), & u > 0, v > 0 \end{cases}$$
For $u < 0$, $v > 0$ $P(x < u) = P(x < u)P(|Y| > -v) = P(x < u)$, with dependent on v .

Therefore, there is no \overline{U} and $\overline{\Sigma}$ that would form a bivariate normal PDF \Rightarrow X and Z' are not bivariate normal.

No. 1, Y ~ N(0; 1), independent \Rightarrow $Z = \begin{bmatrix} x \\ y \end{bmatrix}$ is bivariate normal.

Therefore, $X+Y \sim N([11]\times[0]; [11][0], [11][0]) = N(0; 2)$ $P(X^{2} < u) = P(-\sqrt{u} < X < \sqrt{u}) \cdot I_{u>0} = (2 \cdot P(\sqrt{u}) - 1) \cdot I_{u>0} - \frac{1}{2}$ $f_{x^{2}}(u) = (P(X^{2} < u))_{u}^{1} = I_{u>0} \cdot 2 \cdot P(\sqrt{u}) \cdot \frac{1}{2\sqrt{u}} = I_{u>0} \cdot \frac{1}{2\sqrt{u}} e^{-\frac{1}{2}}$

$$F_{X,Y}(u,v) = P(X^{2} < u, Y^{2} < v) = P(-\sqrt{u} < X < \sqrt{u}, -\sqrt{v} < Y < \sqrt{v}) \cdot I_{u>0} \cdot I_{$$

Therefore,
$$\chi^{2}$$
 and Y^{2} are also independent.
 $f_{\chi^{2}, Y^{2}}(u, v) = f_{\chi^{2}}(u) \cdot f_{Y^{2}}(v) = \frac{1}{2\pi \sqrt{uv'}} e^{-\frac{u}{2} - \frac{v}{2}} \cdot \int_{u>0}^{u} \cdot \int_{v>0}^{t_{v}} dv = \int_{t_{v}}^{t_{v}} \int_{v}^{u} \int_{u=0}^{t_{v}} \int_{v}^{u} \int_{v}$