Stat. 4. UE

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

1. a) $X \sim \mathcal{N}(\mu, \sigma^2) \Rightarrow Z = X \cdot \frac{1}{\sigma} - \frac{\mu}{\sigma} \sim \mathcal{N}(0,1)$.

$$M_X(t) = M_{\sigma Z+\mu}(t) = e^{\mu t} M_Z(\sigma t) \overset{\circledast}{=} e^{\mu t} e^{\frac{1}{2}\sigma^2 t^2}$$

$\circledast$: $M_Z(t) = \mathbb{E}(e^{tX}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} e^{tx}\, dx = \int_{-\infty}^{\infty} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2}}_{f_X(x-t)} e^{\frac{1}{2}t^2}\, dx = e^{\frac{1}{2}t^2}$.

b) $X \sim \mathcal{N}(\mu, \sigma^2)$, $Y = aX + b$. Show that $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$. We must have $a \neq 0$.

Define the mapping $g: \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto ax+b \end{cases}$ with inverse $h: x \mapsto \frac{x-b}{a}$.
h has Jacobian $(\frac{1}{a}) \neq 0$. Therefore, by the change of variable theorem

$$f_Y(y) = f_X(h(y))\, |J(y)| = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(\frac{y-b}{a} - \mu)^2 / 2\sigma^2} \frac{1}{|a|} = \frac{1}{\sqrt{2\pi\sigma^2 a^2}} e^{-(y-(a\mu+b))^2 / 2a^2\sigma^2}.$$

c) Let $X_1, \dots, X_n$ iid, $X_1 \sim \mathcal{N}(\mu, \sigma^2)$. Show that $\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$.

$\circledast$ We know that $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$, $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$ independent $\Rightarrow X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Induction: $n=1$: ✓ $n \mapsto n+1$: Let $\bar{X}^k := \frac{1}{k} \sum_{i=1}^{k} X_k$. We get $\bar{X}^{n+1} = \frac{n\bar{X}^n + X_{n+1}}{n+1} =$

$$= \frac{n}{n+1} \bar{X}^n + \frac{1}{n+1} X_{n+1} \sim \mathcal{N}\left(\underbrace{\left(\frac{1}{n+1} + \frac{n}{n+1}\right)}_{=1} \mu,\ \sigma^2 \underbrace{\left(\left(\frac{1}{1+n}\right)^2 + \frac{1}{n}\left(\frac{n}{n+1}\right)^2 \right)}_{=\frac{1}{n+1}} \right).$$

by induction hypothesis: $\bar{X}^n \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$

with b): $\frac{n}{n+1}\bar{X}^n \sim \mathcal{N}\left(\frac{n}{n+1}\mu, \left(\frac{n}{n+1}\right)^2 \frac{\sigma^2}{n}\right)$

$\frac{1}{n+1} X_{n+1} \sim \mathcal{N}\left(\frac{1}{n+1}\mu, \left(\frac{1}{n+1}\right)^2 \sigma^2\right)$

with $\circledast$

2. a) $X \sim \mathcal{P}(\lambda_1)$, $Y \sim \mathcal{P}(\lambda_2)$ independent. Show that $X + Y \sim \mathcal{P}(\lambda_1 + \lambda_2)$. $\left[\mathbb{P}(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}, k \geq 0\right]$

$$\mathbb{P}(X+Y=k) = \sum_{i=0}^{k} \mathbb{P}(X=i)\,\mathbb{P}(Y=k-i) = \sum_{i=0}^{k} \frac{\lambda_1^i}{i!} e^{-\lambda_1} \frac{\lambda_2^{(k-i)}}{(k-i)!} e^{-\lambda_2} =$$

$$= e^{-(\lambda_1+\lambda_2)} \frac{1}{k!} \sum_{i=0}^{k} \binom{k}{i} \lambda_1^i \lambda_2^{(k-i)} = e^{-(\lambda_1+\lambda_2)} \frac{1}{k!} (\lambda_1 + \lambda_2)^k.$$

$\left[f(x; 2, \lambda) = \frac{\lambda^2 x e^{-\lambda x}}{\underbrace{\Gamma(2)}_{=1}}\right]$

b) Let $U, V \sim \exp(\lambda)$ i.i.d. $\left[f(x;\lambda) = \lambda e^{-\lambda x} \mathbb{1}_{[x \geq 0]}\right]$. Show $\begin{cases} U+V \sim \Gamma(2,\lambda) \\ \min(U,V) \sim \exp(2\lambda). \end{cases}$

- $f_{U+V}(z) = \int_{-\infty}^{\infty} f_U(z-v) f_V(v)\, dv = \int_{-\infty}^{\infty} \lambda e^{-\lambda(z-v)} \mathbb{1}_{[z \geq v]} \lambda e^{-\lambda v} \mathbb{1}_{[v \geq 0]}\, dv =$
  $= \mathbb{1}_{[z \geq 0]} \int_0^z \lambda^2 e^{-\lambda z}\, dv = \lambda^2 e^{-\lambda z} \cdot z \cdot \mathbb{1}_{[z \geq 0]}$.
- $\mathbb{P}(\min U,V \leq z) = 1 - \mathbb{P}(U > z)\,\mathbb{P}(V > z) = 1 - (1 - \mathbb{P}(U \leq z))(1 - \mathbb{P}(V \leq z)) =$
  $= -\mathbb{P}(U \leq z)\,\mathbb{P}(V \leq z) + \mathbb{P}(U \leq z) + \mathbb{P}(V \leq z) = -(1 - e^{-\lambda z})^2 + 2(1 - e^{-\lambda z}) = -1 + \cancel{2e^{-\lambda z}} - e^{-2\lambda z} + 2 - \cancel{2e^{-\lambda z}}$
  $= 1 - e^{-2\lambda z}$, which is the cdf of $\exp(2\lambda)$. (Wlog. $z \geq 0$, else $\mathbb{P}(\min \leq z) = 0$ obviously).

③ A, B, C i.i.d., $A \sim \mathcal{U}(0,1)$.

a) What is the probability that $Ax^2 + Bx + C = 0$ has real roots?

$ax^2 + bx + c = 0$ has real roots $\Leftrightarrow b^2 \geq 4ac \Leftrightarrow b \geq 2\sqrt{ac} \Leftrightarrow c \leq b^2/4a$

$$\lambda^3\left(\left\{\begin{pmatrix}a\\b\\c\end{pmatrix} \in [0,1]^3 : b^2 \geq 4ac\right\}\right) = \int_0^{1/4}\int_0^1\int_{2\sqrt{ac}}^1 db\, dc\, da + \int_{1/4}^1\int_0^1\int_0^{b^2/4a} dc\, db\, da = \circledast,$$

because: Case 1: $a \leq \frac{1}{4}$: $2\sqrt{ac} \leq \sqrt{c} \leq 1 \quad \forall c \in [0,1]$;
Case 2: $a \geq \frac{1}{4}$: $4a \geq 1 \Rightarrow b^2/4a \leq b^2 \leq 1 \quad \forall b \in [0,1]$.

By using a calculator, we get $\circledast = \frac{5}{36} + \frac{\log 2}{6} \approx 25{,}4\,\%$.

b) „runif(n)" returns a vector of n random numbers (uniformly distributed on $[0,1]$).
Moreover, $\sum_{i=1}^{n} b_i = |\{i \in [n] : b_i = \text{TRUE}\}|$ if $b_i \in \{\text{TRUE}, \text{FALSE}\}$ $\forall i \in [n]$.
So the code chooses 10.000 triples (a,b,c) and returns the percentage of triples where $ax^2 + bx + c$ has real roots, so it approximates $\circledast$.

④ a) and c) $\Rightarrow$ see R file

b) $X_1 \sim \mathcal{N}(5, 2^2)$, $X_1, \ldots, X_{50}$ i.i.d., $S := \sum_{i=1}^{50} X_i$, $\overline{X} := \frac{1}{50} S$.

We have $S \sim \mathcal{N}(\mu, \sigma^2)$ where $\mu = \sum_{i=1}^{50} \mu(X_i) = 5 \cdot 50 = 250$
and $\sigma^2 = \sum_{i=1}^{50} \sigma^2(X_i) = 50 \cdot 2^2 = 200$

and $\overline{X} \sim \mathcal{N}(\mu, \sigma^2)$ where $\mu = \mu(X_1) = 5$
and $\sigma^2 = \dfrac{\sigma^2(X_1)}{50} = \dfrac{2^2}{50}$.

⑤ Let $\overline{X}_1, \overline{X}_2$ be the means of two indep. samples of size n with variance σ^2.
Find $n \in \mathbb{N}$ s.t. $\mathbb{P}(|\overline{X}_1 - \overline{X}_2| < \sigma/50) \approx 0{,}99$

- $\mathbb{E}(\overline{X}_1 - \overline{X}_2) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}(X_{1,i} - Y_{1,i})\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_{1,i} - Y_{1,i}) = 0$,
- $\mathbb{V}(\overline{X}_1 - \overline{X}_2) = \mathbb{V}(\overline{X}_1) + \mathbb{V}(\overline{X}_2) = \frac{1}{n^2}\sum_{i=1}^{n}\mathbb{V}(X_{1,i}) + \mathbb{V}(X_{2,i}) = \frac{1}{n^2} 2n\sigma^2 = \frac{2\sigma^2}{n}$.

Chebyshev: $\mathbb{V}(X) < \infty \Rightarrow \forall c > 0$: $\mathbb{P}(|X - \mathbb{E}(X)| \geq c) \leq \dfrac{\mathbb{V}(X)}{c^2}$.

$$\Rightarrow \mathbb{P}(|\overline{X}_1 - \overline{X}_2| < \sigma/50) = 1 - \mathbb{P}(|\overline{X}_1 - \overline{X}_2| \geq \sigma/50) \overset{\mathbb{E}=0}{\geq} 1 - \frac{2\sigma^2}{n}\left(\frac{50}{\sigma}\right)^2 = 1 - \frac{5000}{n} \overset{!}{\geq} 0{,}99$$

$\Rightarrow n - 5000 \geq 0{,}99n \Leftrightarrow 0{,}01n \geq 5000 \Leftrightarrow n \geq \dfrac{5000}{0{,}01} = \underline{\underline{500.000}}$.