(1) **The mean of independent normal distributions**

(a) Show that the moment generating function (mgf) of $X \sim \mathcal{N}(\mu, \sigma^2)$ is of the form

$$M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}.$$

(b) Let $X \sim \mathcal{N}(\mu, \sigma^2)$ and let $Y = aX + b$ with fixed real constants a and b. Show that $Y \sim \mathcal{N}(a\mu + b,\, a^2\sigma^2)$.

(c) Let $X_1, \dots X_n$ be independent identically distributed random variables with $X_1 \sim \mathcal{N}(\mu, \sigma^2)$. Show that the mean $\bar{X} = \frac{1}{n}(X_1 + \dots + X_n)$ is also normally distributed and $\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$.

a) $M_X(t) = \mathbb{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \exp\left(\frac{2\sigma^2 tx - x^2 + 2x\mu - \mu^2}{2\sigma^2}\right) dx$

$\left| \begin{array}{l} u = \frac{x}{\sqrt{2}\sigma} - \left(\frac{\sigma t}{\sqrt{2}} + \frac{\mu}{\sqrt{2}\sigma}\right) \\ \frac{du}{dx} = \frac{1}{\sqrt{2}\sigma} \\ dx = \sqrt{2}\sigma\, du \end{array} \right| = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp\left(-u^2 + \frac{\sigma^2 t^2}{2} + t\mu\right) du = e^{\mu t + \frac{\sigma^2 t^2}{2}}$

$\left(\frac{x}{\sqrt{2}\sigma} - \left(\frac{\sigma t}{\sqrt{2}} + \frac{\mu}{\sqrt{2}\sigma}\right)\right)^2 = \frac{x^2}{2\sigma^2} - 2\frac{x}{\sqrt{2}\sigma}\left(\frac{\sigma t}{\sqrt{2}} + \frac{\mu}{\sqrt{2}\sigma}\right) + \frac{\sigma^2 t^2}{2} + t\mu + \frac{\mu^2}{2\sigma^2}$

b) Case 1: $a = 0$, then $Y = b$, hence $\mathbb{P}(Y \leq c) = \begin{cases} 0, & \text{if } c < b \\ 1, & \text{else} \end{cases}$, could maybe be viewed as the limit „$\mathcal{N}(b, 0)$"

Case 2: $y = ax + b \Leftrightarrow x = \frac{y-b}{a}$ and $\left(\frac{y-b}{a} - \mu\right)^2 = \frac{(y - (\mu a + b))^2}{a^2}$

$f_Y(y) = f_X\left(\frac{y-b}{a}\right)\frac{1}{a} = \frac{1}{\sqrt{2\pi}\sigma a} \exp\left(-\frac{(y - (\mu a + b))^2}{2\sigma^2 a^2}\right)$, hence $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

c) $Y_n := \sum_{i=1}^{n} X_i$

Claim: $\forall n \in \mathbb{N}: Y_n \sim \mathcal{N}(n\mu, n\sigma^2)$

$n = 1$: $Y_1 = X_1 \sim \mathcal{N}(\mu, \sigma^2)$

$n \rightsquigarrow n+1$: Assume that $Y_n \sim \mathcal{N}(n\mu, n\sigma^2)$ for $n \in \mathbb{N}$, then

$Y_{n+1} = Y_n + X_{n+1} \sim \mathcal{N}((n+1)\mu, (n+1)\sigma^2)$, because

$f_{Y_n + X_{n+1}}(z) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(z-t-\mu)^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi n}\sigma} \exp\left(\frac{(t-n\mu)^2}{2n\sigma^2}\right) dt$

$= \frac{1}{2\pi\sigma^2\sqrt{n}} \int_{-\infty}^{\infty} \exp\left(-\left(\frac{z^2}{2\sigma^2} - \frac{zt}{\sigma^2} + \frac{t^2}{2\sigma^2} - \frac{z\mu}{\sigma^2} + \frac{t\mu}{\sigma^2} + \frac{\mu^2}{2\sigma^2} + \frac{t^2}{2n\sigma^2} - \frac{\mu t}{\sigma^2} - \frac{n\mu^2}{2\sigma^2}\right)\right) dt$

$= \frac{\sqrt{2}\sigma(n+1)}{2\pi\sigma^2\sqrt{n}} \int_{-\infty}^{\infty} \exp\left(-u^2 + \frac{z^2(n+1)^2 - 1}{2\sigma^2} + \frac{z\mu}{\sigma^2} + \frac{\mu^2}{2\sigma^2} + \frac{n\mu^2}{2\sigma^2}\right) du$

$= \frac{n+1}{\sqrt{2}\pi\sigma} \exp\left(-\frac{(z-(n+1)\mu)^2}{2n^2\sigma^2}\right) \int_{-\infty}^{\infty} e^{-u^2} du = \frac{n+1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(z-(n+1)\mu)^2}{2n^2\sigma^2}\right)$

$\left(\frac{t}{\sqrt{2}\sigma(n+1)} - \frac{z(n+1)}{\sqrt{2}\sigma}\right)^2 = \frac{t^2}{2\sigma^2(n+1)^2} - \frac{zt}{\sigma^2} + \frac{z^2(n+1)^2}{2\sigma^2}$

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

(2) **Sum of two independent distributions**

(a) Let $X \sim \mathcal{P}(\lambda_1)$ and $Y \sim \mathcal{P}(\lambda_2)$ be two independent Poisson random variables. Show that

$$X + Y \sim \mathcal{P}(\lambda_1 + \lambda_2).$$

(b) Let U and V be two independent random variables with exponential distribution $\exp(\lambda)$. Show that

$$U + V \sim Gamma(2, \lambda) \qquad \text{and}$$
$$\min\{U, V\} \sim \exp(2\lambda).$$

Hint: It is useful to use moment generating functions. Recall, the pdf of a random variable $X \sim Gamma(\alpha, \beta)$ is

$$f(x) = \begin{cases} \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\Gamma(\alpha)\,\beta^\alpha} & x > 0 \\ 0, & x \leq 0 \end{cases},$$

and its mgf is of the form $\left(\frac{1}{1-\beta t}\right)^\alpha$ for $t < \frac{1}{\beta}$. Particularly, the pdf of a random variable $X \sim \exp(\lambda) = Gamma(1, \frac{1}{\lambda})$ is of the form

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x \leq 0 \end{cases}.$$

a) $M_X(t) = \mathbb{E}(e^{tX}) = e^{-\lambda_1} \sum_{k=0}^{\infty} e^{tk} \frac{\lambda_1^k}{k!} = e^{-\lambda_1} \sum_{k=0}^{\infty} \frac{(\lambda_1 e^t)^k}{k!} = e^{-\lambda_1} e^{\lambda_1 e^t} = e^{\lambda_1(e^t - 1)}$

$M_Y(t) = e^{\lambda_2 (e^t - 1)}$, $\quad M_X(t)\, M_Y(t) = e^{(\lambda_1 + \lambda_2)(e^t - 1)} = M_{X+Y}(t)$

Hence, $X + Y \sim \mathcal{P}(\lambda_1 + \lambda_2)$

b) $M_U(t) = M_V(t) = \frac{1}{1 - \lambda t}$, $\quad M_{U+V}(t) = M_U(t) M_V(t) = \frac{1}{(1-\lambda t)^2}$,

Hence, $U + V \sim Gamma(2, \lambda)$, used $\exp(\lambda) = Gamma(1, \lambda)$!!!,

with pdf $f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha - 1} e^{-\lambda x}$, $0 < x < \infty$

$\mathbb{P}(\min(U,V) < z) = 1 - \mathbb{P}(U \geq z \wedge V \geq z) = 1 - \mathbb{P}(U \geq z)\,\mathbb{P}(V \geq z) = 1 - (1 - \mathbb{P}(U < z))^2$

$= 1 - (1 - (1 - e^{-\lambda z}))^2 = 1 - e^{-2\lambda z}$, where $F_U(x) = 1 - e^{-\lambda x}$, $0 < x < \infty$

is the cdf of an exponential distribution.

(3) **Real roots**
Let A, B and C be independent random variables, uniformly distributed on $(0,1)$.

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

(b) Consider the following code in R.
What does it do and how is it related to your solution in part (a)?

```
n=10000
a=runif(n)
b=runif(n)
c=runif(n)
sum(b^2>4*a*c)/n
```

Hint: In HW2/ex. 3(b) we showed that if X has uniform $(0,1)$ distribution then $-\log X$ has exponential distribution $\exp(1)$. In an analogue way, one can prove that $-s\log X \sim \exp(\frac{1}{s})$ for any $s > 0$. Also, in HW4/ex. 2(b) we proved that the sum of two independent exponential distributions is a gamma distribution. Namely, if $X \sim \exp(1)$ and $Y \sim \exp(1)$ are independent then $X + Y \sim Gamma(2,1)$.

a) $Ax^2+Bx+C=0 \Leftrightarrow x=\frac{-B\pm\sqrt{B^2-4AC}}{2A}$, hence the quadratic equation has real roots if and only if $B^2-4AC\geq 0$.

$$\mathbb{P}(B^2-4AC\geq 0)=\mathbb{P}(B^2\geq 4AC)=\mathbb{P}(\log(B^2)\geq\log(4AC))=\mathbb{P}(\log(B^2)\geq\log(4)+\log(AC))$$
$$=\mathbb{P}(\log(B^2)-\log(AC)\geq\log(4))=\mathbb{P}(-\log(AC)\geq-\log(B^2)+\log(4))$$

$-\log(A)\sim\exp(1)$, $-\log(C)\sim\exp(1)$, $-\log(AC)=(-\log(A))+(-\log(C))\sim Gamma(2,1)$

$g:(0,1)\to(0,\infty): x\mapsto-\log(x^2)$ has got the inverse $h:(0,\infty)\to(0,1): y\mapsto e^{-\frac{y}{2}}$

$f_B(x)=\mathbb{1}_{(0,1)}(x) \Rightarrow f_{-\log(B^2)}(x)=\mathbb{1}_{(0,1)}(e^{-\frac{x}{2}})\, e^{-\frac{x}{2}}=\mathbb{1}_{(0,\infty)}(x)\,\frac{1}{2}e^{-\frac{x}{2}}$, hence $-\log(B^2)\sim\exp(\frac{1}{2})$

Clearly, $-\log(B^2)$ and $-\log(AC)$ are independent and we conclude

$$\mathbb{P}(-\log(AC)\geq-\log(B^2)+\log(4))=\int_{\log(4)}^{\infty}\int_0^{y-\log(4)} f_{-\log(AC)}(y)\, f_{-\log(B^2)}(x)\,dx\,dy$$
$$=\int_{\log(4)}^{\infty} ye^{-y}\int_0^{y-\log(4)}\frac{1}{2}e^{-\frac{x}{2}}\,dx\,dy$$
$$=\int_{\log(4)}^{\infty} ye^{-y}\left(1-\exp\left(\frac{\log(4)-y}{2}\right)\right)dy$$
$$=\int_{\log(4)}^{\infty} ye^{-y}\,dy-\int_{\log(4)}^{\infty}2ye^{-\frac{3y}{2}}\,dy$$
$$=\frac{1}{4}(1+\log(4))-\frac{1}{9}(1+\log(8))\approx 0{,}2544$$

b) The code gives an approximation of the probability that was calculated in (a).

(4) **Sum and average**

Let X be a random variable with $\mathcal{N}(5, 2^2)$. Let $X_1, X_2, \ldots, X_{50}$ be independent identically distributed copies of X. Let S be their sum and $\bar{X}$ their average, i.e.

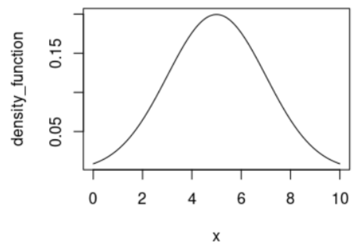
$$S = X_1 + \cdots + X_{50} \qquad \text{and} \qquad \bar{X} = \frac{1}{50}(X_1 + \cdots + X_{50}).$$

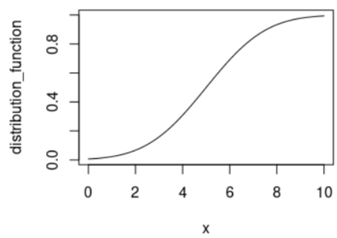
(a) Plot the density and the distribution function for X using R.

(b) What are the expectation and the standard deviation of S and of $\bar{X}$?

(c) Generate a sample of 50 numbers from $\mathcal{N}(5, 2^2)$. Plot the histogram for this sample. Do the same for a sample of 500 numbers from $\mathcal{N}(5, 2^2)$.

a)

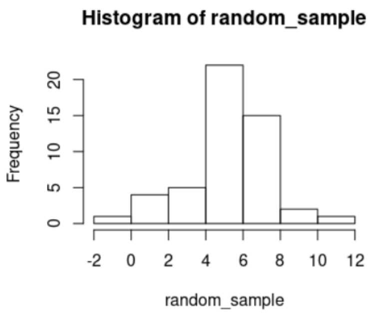




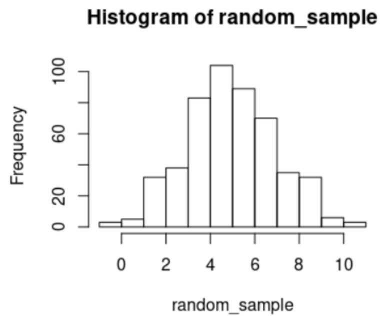
b) By problem 1 we have $\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$ and $S = n\bar{X} \sim \mathcal{N}(n\mu, n\sigma^2)$

Hence $\mathbb{E}(\bar{X}) = \mu$, $\sqrt{\mathbb{V}\text{ar}(\bar{x})} = \frac{\sigma}{\sqrt{n}}$, $\mathbb{E}(S) = n\mu$, $\sqrt{\mathbb{V}\text{ar}(S)} = \sqrt{n}\,\sigma$

c) $n=50$



$n=500$



(5) **Central Limit Theorem**
Let $\bar{X}_1$ and $\bar{X}_2$ be the means of two independent samples of size n from the same population with variance σ^2. Use the Central limit theorem to find a value for n so that

$$P(|\bar{X}_1 - \bar{X}_2| < \frac{\sigma}{50}) \approx 0.99.$$

Justify your calculations.

We rename $\overline{Y_n} := \overline{X_1}$ and $\overline{Z_n} := \overline{X_2}$

$$\mathbb{E}(\overline{Y_n} - \overline{Z_n}) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} Y_i - \frac{1}{n}\sum_{i=1}^{n} Z_i\right) = \frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{E}(Y_i) - \mathbb{E}(Z_i)\right) = 0$$

$$\mathbb{V}(\overline{Y_n} - \overline{Z_n}) = \mathbb{V}(\overline{Y_n}) + \mathbb{V}(-\overline{Z_n}) = \frac{1}{n^2}\sum_{i=1}^{n}\left(\mathbb{V}(Y_i) + \mathbb{V}(Z_i)\right) = \frac{2\sigma^2}{n}$$

Chebyshev's inequality (13.14 in Kusolitsch) says that

$$\mathbb{P}(|(\overline{Y_n} - \overline{Z_n}) - \underbrace{\mu}_{=0}| \geq \frac{\sigma}{50}) \leq \frac{\mathbb{V}(\overline{Y_n} - \overline{Z_n})}{\left(\frac{\sigma}{50}\right)^2} = \frac{\left(\frac{2\sigma}{n}\right)^2}{\left(\frac{\sigma}{50}\right)^2} = \frac{100^2}{n^2}, \text{ hence}$$

$$\mathbb{P}(|\overline{Y_n} - \overline{Z_n}| < \frac{\sigma}{50}) \geq 1 - \frac{100^2}{n^2} \stackrel{!}{=} \frac{99}{100} \Leftrightarrow \frac{1}{100} = \frac{100^2}{n^2} \Leftrightarrow n^2 = 100^3 \Leftrightarrow n = 100^{\frac{3}{2}} = \underline{\underline{1000}}$$