# Density Transformations & random number generation

## Exercise T4.1: From Pseudo-random number generators to density transformation (tutorial)

(a) What is the ICA Problem?

(b) What is the cumulative density function (cdf) to a probability density function (pdf)?

(c) How do you transform densities while conserving probabilities?

## Exercise H4.1: The Inverse CDF method (homework, 4 points)

**Background:** If $F_X(x)$ is the cumulative distribution function (cdf) of a random variable $X$, then the random variable $Z = F_X(X)$ is uniformly distributed on the interval $[0, 1]$. This result provides a general recipe to generate samples $\tilde{x}$ of a random variable $X$ with a desired probability density function (pdf) $p_X(x)$ from uniformly distributed random numbers $\tilde{z} \in [0, 1]$:

1. Compute the cdf $F_X(x)$ of the desired pdf $p_X(x)$
2. Determine the inverse transformation $F^{-1}$.
3. Sample uniformly distributed numbers $\tilde{z}$ in $[0, 1]$.
4. Get the samples $\tilde{x} = F^{-1}(\tilde{z})$ from $X$.

The pdf of a Laplace distribution with location parameter $\mu$ (= mean), and scale parameter $b > 0$ (variance = $2b^2$) is given by

$$p_X(x) = \frac{1}{2b} \exp\left(-\frac{|x-\mu|}{b}\right).$$

**Task:**

(a) Following the procedure above, derive a formula to generate samples of a scalar random variable with a Laplacian distribution from uniformly distributed random numbers.

<u>**Solution**</u>

The pdf of the Laplacian distribution is given as:

$$p(x) = \frac{1}{2b} \begin{cases} \exp\left(-\frac{\mu-x}{b}\right) & x < \mu \\ \exp\left(-\frac{x-\mu}{b}\right) & x \geq \mu \end{cases}$$

The cumulative distribution function $F(x) = P(X \leq x) = \int_{-\infty}^{x} p(y)dy$ reads

$$F(x) = \frac{1}{2b} \begin{cases} \int_{-\infty}^{x} \exp\left(-\frac{\mu-y}{b}\right) dy & x < \mu \\ \int_{-\infty}^{\mu} \exp\left(-\frac{\mu-y}{b}\right) dy + \int_{\mu}^{x} \exp\left(-\frac{y-\mu}{b}\right) dy & x \geq \mu \end{cases}$$

i.e.

$$F(x) = \begin{cases} \frac{1}{2b}\, b\ \exp\left(\frac{y-\mu}{b}\right) \Big|_{y=-\infty}^{y=x} & x < \mu \\ \frac{1}{2b}\, b\ \exp\left(\frac{y-\mu}{b}\right) \Big|_{y=-\infty}^{y=\mu} - \frac{1}{2b}\, b\ \exp\left(\frac{\mu-y}{b}\right) \Big|_{y=\mu}^{y=x} & x \geq \mu \end{cases}$$

resulting in

$$F(x) = \begin{cases} \frac{1}{2} e^{\frac{x-\mu}{b}} & & x < \mu \\ \frac{1}{2} - \frac{1}{2}(e^{\frac{\mu-x}{b}} - 1) & = \quad 1 - \frac{1}{2} e^{\frac{\mu-x}{b}} & x \geq \mu \end{cases}$$

For the inverse cdf $F^{-1}(z)$ we get for

- $x < \mu$, with $z = \frac{1}{2} e^{\frac{x-\mu}{b}} < \frac{1}{2}$

$$b \ln(2z) = x - \mu \qquad \rightarrow \qquad F^{-1}(z) = x = b \ln(2z) + \mu$$

- $x \geq \mu$ with $z = 1 - \frac{1}{2} e^{\frac{\mu-x}{b}} \geq \frac{1}{2}$

$$b \ln(2 - 2z) = \mu - x \qquad \rightarrow \qquad F^{-1}(z) = x = -b \ln(2 - 2z) + \mu$$

Taking $y := z - \frac{1}{2}$ allows us to write the above in a more compact form:

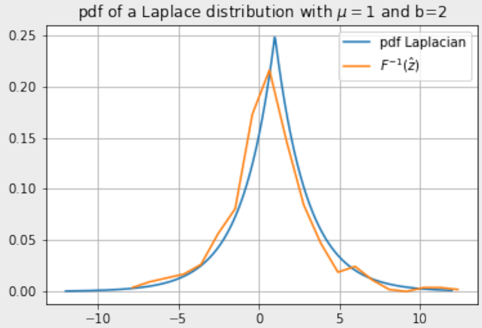
$$\begin{aligned} x = \left\{ \begin{array}{ll} +b \ln(1 + 2y) + \mu & y < 0 \\ -b \ln(1 - 2y) + \mu & y \geq 0 \end{array} \right\} &= -\ \mathrm{sign}(y) b \ln(1 - 2|y|) + \mu \\ &= \mathrm{sign}(y) b \ln(2|y|) + \mu \\ &\overset{y=z-\frac{1}{2}}{=} \mathrm{sign}\left(z - \frac{1}{2}\right) b \ln\left(2\left|z - \frac{1}{2}\right|\right) + \mu \end{aligned}$$

With this formula we get samples of a random variable with a Laplacian distribution using samples from uniformly distributed variable $z$.

(b) Implement your procedure for verification and generate 500 samples for a Laplacian random variable $X$ with a specific mean $\mu = 1$ and scale parameter $b = 2$.

(c) Plot a density estimate (e.g. normalized histogram) for these samples overlayed with the pdf $p_X(x)$ from above.

**Solution**


pdf of a Laplace distribution with $\mu = 1$ and b=2
pdf Laplacian
$F^{-1}(\hat{z})$
0.25
0.20
0.15
0.10
0.05
0.00
−10
−5
0
5
10

## Exercise H4.2: Density Transformations (homework, 6 points)

**Background:** Let $f(\underline{\mathbf{x}}) = f(x_1, \ldots, x_N)$ be a function of $\underline{\mathbf{x}} \in \Omega \subset \mathbb{R}^N$ and $\underline{\mathbf{g}} : \underline{\mathbf{x}} \mapsto \underline{\mathbf{g}}(\underline{\mathbf{x}})$ be a mapping with which we change the variables $\underline{\mathbf{x}}$ to a new coordinate system with coordinates $\underline{\mathbf{u}} = (g_1(\underline{\mathbf{x}}), \ldots, g_N(\underline{\mathbf{x}})) = (u_1, \ldots, u_N)^\top$, whose inverse mapping $\underline{\mathbf{g}}^{-1}(\underline{\mathbf{u}}) = \underline{\mathbf{x}}$ exists and is differentiable.
As we change the coordinate system, the integral over $f(\cdot)$ changes according to

$$\int_\Omega f(\underline{\mathbf{x}}) \mathbf{d}\underline{\mathbf{x}} = \int_{g(\Omega)} f(\underbrace{\underline{\mathbf{g}}^{-1}(\underline{\mathbf{u}})}_{=\underline{\mathbf{x}}}) \left| \det \frac{\partial \underline{\mathbf{x}}}{\partial \underline{\mathbf{u}}} \right| \mathbf{d}\underline{\mathbf{u}} = \int_{g(\Omega)} f(\underline{\mathbf{g}}^{-1}(\underline{\mathbf{u}})) \frac{1}{\left| \det \frac{\partial \underline{\mathbf{u}}}{\partial \underline{\mathbf{x}}} \right|} \mathbf{d}\underline{\mathbf{u}},$$

where $\frac{\partial \underline{\mathbf{x}}}{\partial \underline{\mathbf{u}}}$ is the *Jacobi* matrix, which is the matrix of the partial derivatives

$$\frac{\partial \underline{\mathbf{x}}}{\partial \underline{\mathbf{u}}} = \frac{\partial \underline{\mathbf{g}}^{-1}(\underline{\mathbf{u}})}{\partial \underline{\mathbf{u}}} = \begin{pmatrix} \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_1}{\partial u_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_N}{\partial u_1} & \cdots & \frac{\partial x_N}{\partial u_N} \end{pmatrix}$$

and whose determinant $\det \frac{\partial \underline{\mathbf{x}}}{\partial \underline{\mathbf{u}}} = \left( \det \frac{\partial \underline{\mathbf{u}}}{\partial \underline{\mathbf{x}}} \right)^{-1}$ is called the *Jacobi determinant* (also *functional determinant*).

*Remark:* The absolute value of the Jacobi determinant at a point $\underline{\mathbf{u}}_0$ corresponds to the factor by which the inverse mapping function $\underline{\mathbf{g}}^{-1}(\underline{\mathbf{u}})$ expands or shrinks volumes near $\underline{\mathbf{u}}_0$.

Implication: If $f(\underline{\mathbf{x}})$ is the probability density function (pdf) of the $N$-dimensional random vector $(X_1, \ldots, X_N)^\top$ then $f(\underline{\mathbf{g}}^{-1}(\underline{\mathbf{u}})) \left| \det \frac{\partial \underline{\mathbf{x}}}{\partial \underline{\mathbf{u}}} \right|$ is the pdf of the random vector $(U_1, \ldots, U_N)^\top$.

**Task:**

(a) (1 point) Consider the density of a random variable $X$ to be $p_X(x) = e^{-x},\ x \geq 0$.
For the change of variables $u = e^{-x}$ calculate the density $p_U(u)$ of the random variable $U$.

(b) (4 points) Consider two independent and uniformly in the interval $[0, 1]$ distributed random variables $(X_1, X_2)^\top$. The pdf is given by $p_{X_1,X_2}(x_1, x_2) = 1$ in $[0, 1]^2$ and zero otherwise.
Consider the variable transformation $(X_1, X_2)^\top \to (U_1, U_2)$ with
$u_1 = \sqrt{-2 \ln x_1} \cos(2\pi x_2)$ and
$u_2 = \sqrt{-2 \ln x_1} \sin(2\pi x_2)$.
Show that $(U_1, U_2)^\top$ corresponds to two independent unit-variance zero-mean normally distributed random variables.
*Remark:*
This procedure to produce Gaussian samples from uniform random numbers is called the *Box-Muller method*.

(c) (1 point) **Outline** how to generalize the previous result to $N$ dimensions[1], i.e., how to generate samples from a multidimensional Gaussian distribution with mean vector $\underline{\mu}$ and co-

[1] It might help to think of $N$ as even.

variance matrix $\underline{\mathbf{\Sigma}}$ just from uniformly distributed random numbers in $[0,1]^N$. Use the following:

- Any symmetric positive semidefinite matrix (such as the covariance matrix $\underline{\mathbf{\Sigma}}$) has a Cholesky decomposition $\underline{\mathbf{\Sigma}} = \underline{\mathbf{L}}\,\underline{\mathbf{L}}^\top$ (and that can be easily computed numerically).
- If $\underline{\mathbf{L}}$ is a constant matrix and $\underline{\mathbf{X}}$ a random vector then $\mathrm{Cov}(\underline{\mathbf{L}}\,\underline{\mathbf{X}}) = \underline{\mathbf{L}}\,\mathrm{Cov}(\underline{\mathbf{X}})\,\underline{\mathbf{L}}^\top$.
- The covariance matrix of independent unit-variance Gaussian variables is identity, i.e., $\mathrm{Cov}(\underline{\mathbf{X}}) = \underline{\mathbf{I}}$.

Confirm that the above properties hold for your solution (a detailed proof is <u>not</u> necessary).

**<u>Solution</u>**

**(a)** Due to conservation of probability,

$$\int_\Omega p_X(x)dx \quad = \quad \int_{u(\Omega)} p_X(x(u)) \left|\frac{\partial x(u)}{\partial u}\right| du \quad = \quad \int_{u(\Omega)} p_U(u)du \quad = \quad 1$$

with

$$u = e^{-x} \quad \rightarrow \quad x = -\ln(u) \qquad \Omega = [0,\infty) \quad \rightarrow \quad u(\Omega) = [1,0)$$

The Jacobi Determinant in this case is:

$$\left|\frac{dx(u)}{du}\right| = -\frac{1}{u}$$

and therefore:

$$\int_0^\infty e^{-x}dx \quad = \quad \int_1^0 e^{-(-\ln u)}\,\frac{-1}{u}\,du \quad = \quad \int_1^0 -1\,du \quad = \quad \int_0^1 1\,du$$

i.e.

$$p_U(u) = 1$$

**(b)** Apply the density transformation theorem from above. Due to conservation of probability,

$$\int_\Omega p_X(x)dx \quad = \quad \int_{u(\Omega)} p_X(x(u)) \left|\frac{\partial x(u)}{\partial u}\right| du \quad = \quad \int_{u(\Omega)} p_U(u)\,du \quad = \quad 1$$

with

$$u_1 = \sqrt{-2\ln x_1}\cdot\cos(2\pi x_2) \quad \text{and} \quad u_2 = \sqrt{-2\ln x_1}\cdot\sin(2\pi x_2)$$

Knowing that $\det\left(\frac{\partial \underline{\mathbf{x}}}{\partial \underline{\mathbf{u}}}\right) = \frac{1}{\det\left(\frac{\partial \underline{\mathbf{u}}}{\partial \underline{\mathbf{x}}}\right)}$ we first construct the Jacobian:

$$\frac{\partial \underline{\mathbf{u}}}{\partial \underline{\mathbf{x}}} = \begin{pmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} \frac{-\cos(2\pi x_2)}{x_1\sqrt{-2\ln x_1}} & -2\pi\sqrt{-2\ln x_1}\,\sin(2\pi x_2) \\ \frac{-\sin(2\pi x_2)}{x_1\sqrt{-2\ln x_1}} & 2\pi\sqrt{-2\ln x_1}\,\cos(2\pi x_2) \end{pmatrix}$$

followed by its determinant:

$$
\begin{aligned}
\det\left(\frac{\partial \underline{\mathbf{u}}}{\partial \underline{\mathbf{x}}}\right) &= \frac{-\cos(2\pi x_2)}{x_1\cancel{\sqrt{-2\ln x_1}}} \cdot 2\pi\cancel{\sqrt{-2\ln x_1}}\cos(2\pi x_2) \\
&\qquad - \frac{-\sin(2\pi x_2)}{x_1\cancel{\sqrt{-2\ln x_1}}} \cdot (-2)\pi\cancel{\sqrt{-2\ln x_1}}\sin(2\pi x_2) \\
&= \frac{-2\pi\cos^2(2\pi x_2)}{x_1} - \frac{-(-2)\pi\sin^2(2\pi x_2)}{x_1} \\
&= \frac{-2\pi}{x_1}\big(\underbrace{\cos^2(2\pi x_2) + \sin^2(2\pi x_2)}_{=1}\big) \quad = \frac{-2\pi}{x_1}
\end{aligned}
$$

$$
\left|\det\left(\frac{\partial \underline{\mathbf{x}}}{\partial \underline{\mathbf{u}}}\right)\right| = \frac{x_1}{2\pi}
$$

We proceed by expressing $x_1$ in terms of $u_1$ and $u_2$ only:

$$
\begin{aligned}
u_1^2 + u_2^2 &= -2\ln x_1\cos^2(2\pi x_2) - 2\ln x_1\sin^2(2\pi x_2) \\
&= -2\ln(x_1) \\
-\left(\frac{u_1^2 + u_1^2}{2}\right) &= \ln x_1 \\
x_1 &= e^{\left(-\frac{u_1^2+u_2^2}{2}\right)} \\
&= \exp\left(-\frac{u_1^2}{2}\right)\exp\left(-\frac{u_2^2}{2}\right)
\end{aligned}
$$

with the pdf of a Gaussian distributed random variable defined as $\frac{1}{\sqrt{2\pi\sigma^2}}e^{\left(-\frac{z-\mu}{2\sigma^2}\right)}$ and with $\mu = 0$ and $\sigma = 1$, the pdf of the standard normal distribution becomes $\frac{1}{\sqrt{2\pi}}e^{\left(-\frac{z}{2}\right)}$,

$$
\begin{aligned}
p_{U_1,U_2}(\underline{\mathbf{u}}) &= \overbrace{p_{X_1,X_2}(\underline{\mathbf{x}})}^{=1}\left|\det\left(\frac{\partial \underline{\mathbf{x}}}{\partial \underline{\mathbf{u}}}\right)\right| \\
&= 1 \cdot \frac{x_1}{2\pi} = \frac{1}{2\pi} \cdot \exp\left(-\frac{u_1^2}{2}\right) \cdot \exp\left(-\frac{u_2^2}{2}\right) \\
&= \underbrace{\frac{1}{\sqrt{2\pi}}e^{\left(-\frac{u_1^2}{2}\right)}}_{p_{U_1}(u_1)\sim N(0,1)} \cdot \underbrace{\frac{1}{\sqrt{2\pi}}e^{\left(-\frac{u_2^2}{2}\right)}}_{p_{U_2}(u_2)\sim N(0,1)}
\end{aligned}
$$

and since $p_{U_1,U_2}(\underline{\mathbf{u}}) = p_{U_1}(u_1) \cdot p_{U_2}(u_2)$, the random variables $U_1$ and $U_2$ are indepndent.

**(c)** Let $\underline{\mathbf{Z}} = \underline{\mathbf{L}}\,\underline{\mathbf{X}} + \underline{\mu}$ with $\underline{\mathbf{X}}$ being a random vector of independent standard normal distributed numbers. Then $\underline{\mathbf{Z}}$ is normally distributed with covariance matrix $\underline{\mathbf{L}}\,\underline{\mathbf{L}}^\top$ and mean vector $\underline{\mu}$ from the properties above. I.e. sample using Box-Muller method $N$ samples, multiply them from the left side with $\underline{\mathbf{L}}$ and add $\underline{\mu}$.

To elaborate on the solution for **(c)**:

Our solution for **(b)** demonstrates that a pair of transformed variables $u_1$ and $u_2$ correspond to two independent normally distributed random variables. We start by defining multiple pairs of variables $u_i$ and $u_{i+1}$:

$$
\begin{aligned}
u_i &:= \sqrt{-2\ln x_i}\,\cos(2\pi x_{i+1}) \\
u_{i+1} &:= \sqrt{-2\ln x_i}\,\sin(2\pi x_{i+1})
\end{aligned}
$$

for $i \in \{1, 3, 5, \ldots, N\}$
$u_i \sim \mathcal{N}(0,1)$

We now need to show that independence holds after we concatenate all pairs into one random vector $\underline{\mathbf{U}}$. Let $\underline{\mathbf{Z}} = \underline{\mathbf{L}}\,\underline{\mathbf{U}} + \underline{\mu}$ and by applying the properties above:

$$
\begin{aligned}
\mathrm{Cov}(\underline{\mathbf{L}}\,\underline{\mathbf{U}}) &= \underline{\mathbf{L}}\,\mathrm{Cov}(\underline{\mathbf{U}})\underline{\mathbf{L}}^\top \\
&= \underline{\mathbf{L}} \quad \underline{\mathbf{I}} \quad \underline{\mathbf{L}}^\top \\
&= \underline{\mathbf{L}}\,\underline{\mathbf{L}}^\top \\
&= \underline{\mathbf{\Sigma}}
\end{aligned}
$$

**Total 10 points.**