Computational statistics, lecture math 2

Frank Miller, Department of Computer and Information Science, Linköping University

February 2, 2026

LINKÖPING UNIVERSITY

Topics today

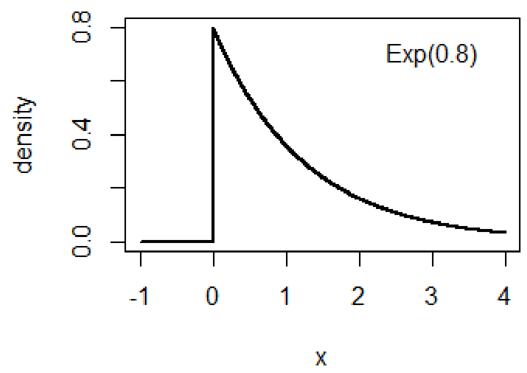
- Probability distributions (reading: GH or Gentle, chapter 1.3) and integration
- Multivariate normal distribution
- Markov chains (reading: GH, chapter 1.7)

Exponential distribution

- The density of the exponential distribution $Exp(\lambda)$ is

$$f(x) = \lambda \exp(-\lambda x)\, \mathbf{1}\{x \geq 0\} = \begin{cases} \lambda \exp(-\lambda x), & \text{if } x \geq 0, \\ 0, & \text{if } x < 0 \end{cases}$$

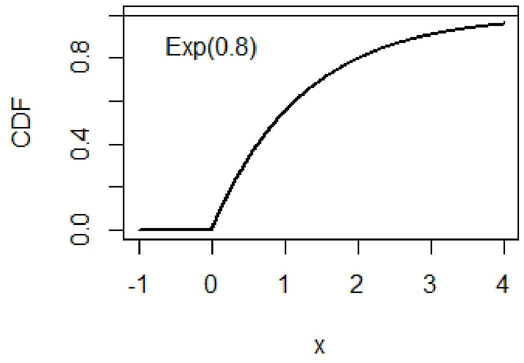
`dexp(x, rate=0.8)`



- Compute the cumulative distribution function (CDF):

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt = \begin{cases} 1 - \exp(-\lambda x), & \text{if } x \geq 0, \\ 0, & \text{if } x < 0 \end{cases}$$

`pexp(x, rate=0.8)`

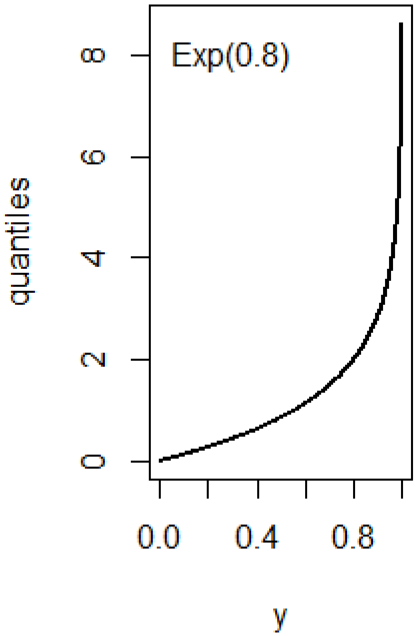


- The inverse cumulative distribution function is $F^{-1}(y) = -\log(1-y)/\lambda$

since $y = 1 - \exp(-\lambda x) \Leftrightarrow \exp(-\lambda x) = 1 - y \;\Leftrightarrow$

$$-\lambda x = \log(1-y) \Rightarrow x = -\log(1-y)/\lambda$$

`qexp(x, rate=0.8)`



Expected value and variance of a distribution

- The probability for $X \in [a, b]$ can be written as integral according to

$$\mathrm{P}(X \in [a, b]) = \int_a^b f(x)dx = \mathrm{F(b)} - \mathrm{F(a)}$$

- The expected value of a distribution with density $f(x)$ is:

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x\, f(x)dx$$

- The variance is

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x - EX)^2\, f(x)dx$$

- What is $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ for the exponential distribution?

Integration

- The expected value of the exponential distribution is:
  $\mathrm{E}(X) = \int_{-\infty}^{\infty} x\, f(x) dx = \int_{-\infty}^{\infty} x\lambda \exp(-\lambda x)\, \mathbf{1}\{x \geq 0\} dx =$

Poisson distribution

- The density of the Poisson distribution $Po(\lambda)$ is

$$f(x) = \frac{\lambda^x}{x!}\exp(-\lambda) \text{ for } x = 0, 1, 2, \ldots$$

```
dpois(x, lambda=2)
```

- The cumulative distribution function is:

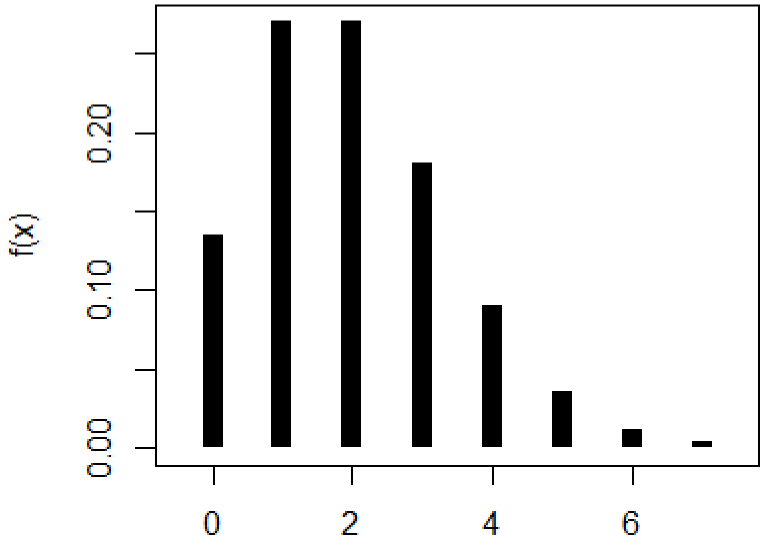
$$F(x) = P(X \leq x) = \sum_{j=0}^{x} f(j) = \exp(-\lambda)\sum_{j=0}^{x}\frac{\lambda^j}{j!}$$

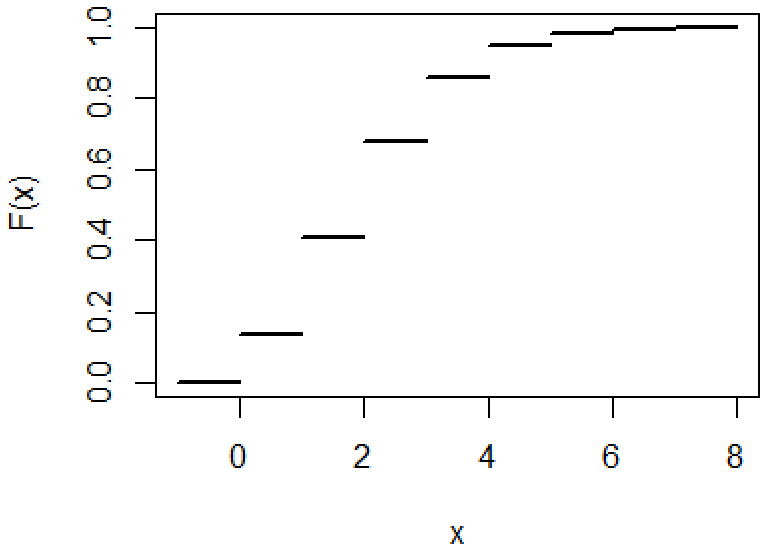
```
ppois(x, lambda=2)
```

- The inverse CDF (quantile function) can be defined as

$$F^{-1}(y) = \min\{x : F(x) \geq y\}$$

```
qpois(x, lambda=2)
```





Exponential family (GH or Gentle, chapter 1.3)

- Exponential distribution: $f(x) = \lambda \exp(-\lambda x)\, \mathbf{1}\{x \geq 0\}$

- Poisson distribution $f(x) = \frac{\lambda^x}{x!} \exp(-\lambda)$

- Exponential, Poisson, normal, ... distributions belong to the exponential family of distributions
- Density of exponential family distribution:

$$f(x) = c_1(x) c_2(\gamma) \exp\left(\sum_{i=1}^{k} y_i(x) \theta_i(\gamma)\right)$$

Multivariate normal distribution

- A multivariate normal distribution $N(\mu, \Sigma)$ with mean vector μ and covariance matrix Σ has density

$$f(\boldsymbol{x}) = \frac{1}{\left((2\pi)^p|\Sigma|\right)^{1/2}} e^{-(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})/2}$$

- If X is a p-dimensional vector of independent standard normal variables (i.e., X is multivariate normal $N(0, I)$, $X \sim N(0, I)$), then $Y = A^T X + \mu$ has a $N(\mu, \Sigma)$-distribution with $\Sigma = A^T A$
- If Σ is given, A with $\Sigma = A^T A$ can be computed, e.g., with the Cholesky decomposition)

Generate a bivariate normal distribution

- Example: Let Σ be the covariance matrix of a bivariate normal distribution with each component having variance 1 and correlation being $\rho, -1 \leq \rho \leq 1$, i.e., $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$
- One can show that $\Sigma = A^T A$ with $A = \begin{pmatrix} 1 & \rho \\ 0 & \sqrt{1-\rho^2} \end{pmatrix}$
- Therefore:
  - Let $X_1 \sim N(0,1);\ X_2 \sim N(0,1)$ be independent, $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$
  - Let $Y = A^T X$, which means:
    $Y_1 = X_1,$
    $Y_2 = \rho X_1 + \sqrt{1-\rho^2} X_2.$
  - Then, $\mu + Y \sim N(\mu, \Sigma)$

Generate multivariate normal distribution in R

- Package **mvtnorm** supplies (besides densities, probabilities, and quantile functions) random generator **rmvnorm** for multivariate normal:

$$\Sigma = \begin{pmatrix} 1 & 0.5 & 0.25 \\ 0.5 & 1 & 0.5 \\ 0.25 & 0.5 & 1 \end{pmatrix}$$

```
library(mvtnorm)
Sig <- matrix(c(1,0.5,0.25, 0.5,1,0.5, 0.25,0.5,1), nrow = 3)
rmvnorm(5, mean = c(2, -1, 0), sigma = Sig)
         [,1]       [,2]        [,3]
[1,] 2.447613 -1.7390355 -0.43272668
[2,] 1.933296 -1.9893701  0.47161126
[3,] 2.872344 -0.6396215  0.61383902
[4,] 2.505797 -0.8080744 -0.51696135
[5,] 2.614221 -1.9280517  0.04471815
```

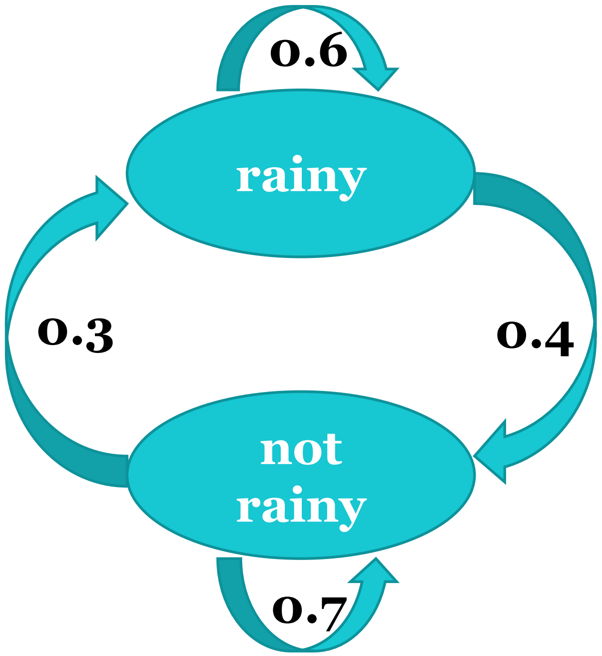
Each row is one generated random vector

- Optionally, method of decomposition can be chosen
- Multivariate t distribution (same package): **rmvt**

Markov chains (reading: GH chapter 1.7)

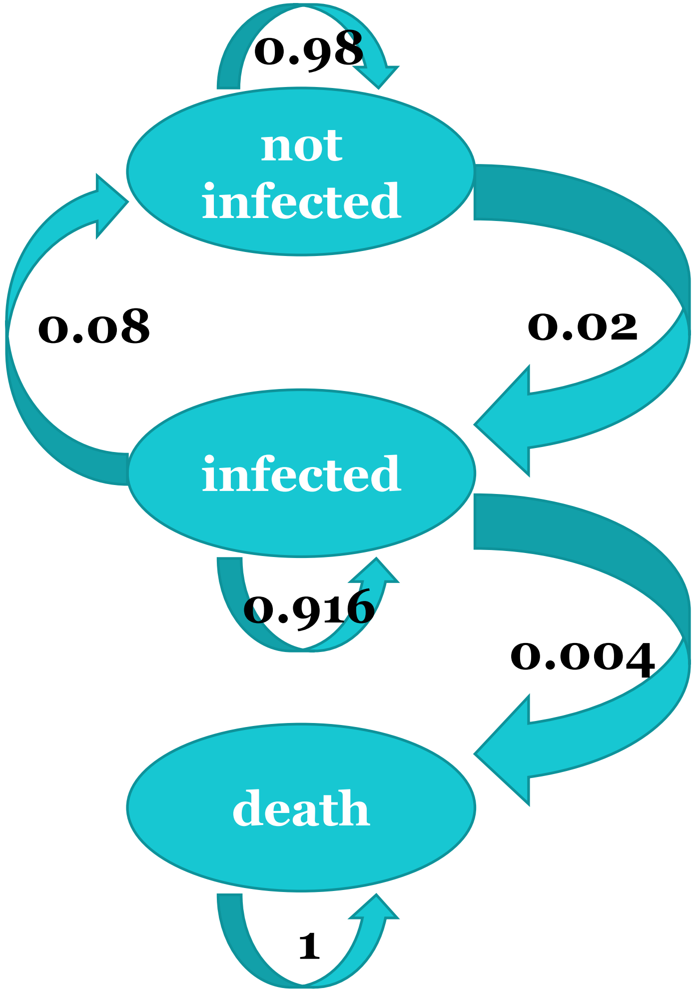
- Consider a sequence of random variables $\left(X^{(t)}\right), t = 0, 1, \ldots$
- Each $X^{(t)}$ can have values in so-called *state space* S (can be discrete; we will later use a continuous one- or multidimensional S)
- A general random sequence (process) is described by specification of $P(X^{(t)}|X^{(t-1)}, \ldots, X^{(0)})$ for all t

- A **Markov chain** is a specific process with *Markov property*
$$\boldsymbol{P\left(X^{(t)}\middle|X^{(t-1)}, \ldots, X^{(0)}\right) = P(X^{(t)}|X^{(t-1)})}$$
- The state of $X^{(t)}$ depends only on the state before and not earlier history ($X^{(t)}$ is memoryless)
- $P(X^{(t)}|X^{(t-1)})$ is called *transition probability*



state space
S={rainy, not rainy}

numbers are
transition probabilities

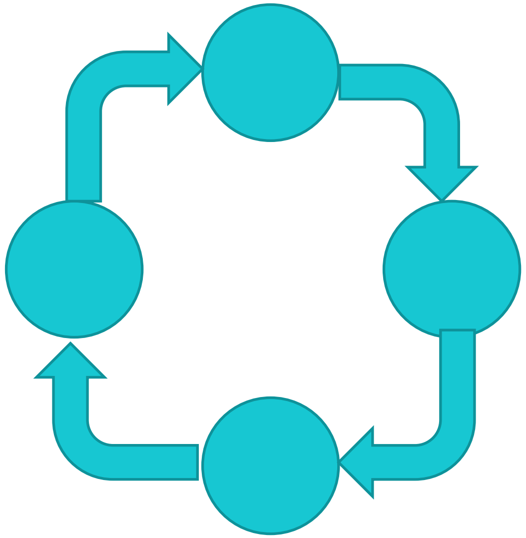


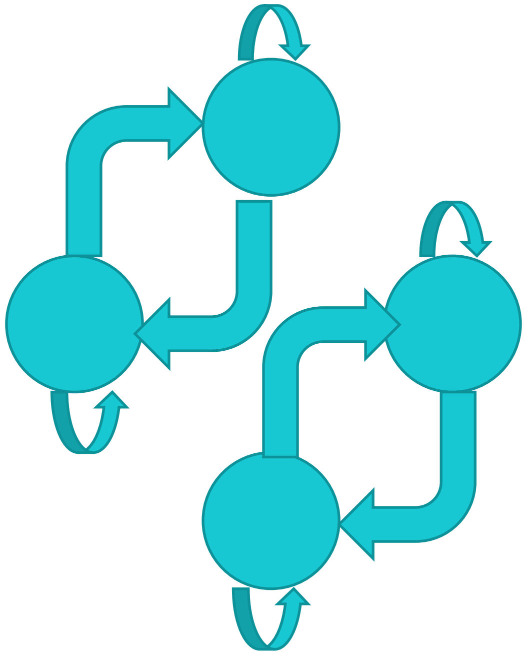
S={not infected, infected, death}

Markov chains

- If S is discrete, we can write $p_{ij}^{(t)} = P(X^{(t)} = j | X^{(t-1)} = i)$
- If S is continuous, $P(X^{(t)} | X^{(t-1)})$ is represented by a cumulative distribution function or density

- A Markov chain is *(time-)homogenous* if
  distribution of $(X^{(t)} | X^{(t-1)})$ is equal for all t
- A Markov chain is *irreducible* if every state in S can be reached from any state
- A state has *period* k if multiples of k steps are necessary to return to it
- A Markov chain is *aperiodic* if each state has period 1





- A Markov chain is *irreducible* if every state in S can be reached from any state
- A state has *period* k if multiples of k steps are necessary to return to it
- A Markov chain is *aperiodic* if each state has period 1