4 Continuous Random Variables

(4.1) Definition. Let Ω be a sample space and $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ a set of events with a probability measure $\Pr : \mathcal{A} \to \mathbb{R}$. A mapping $X : \Omega \to \mathbb{R}$ is called a *random variable* if

$$X^{-1}((-\infty, x]) \ = \ \{\omega \in \Omega \mid X(\omega) \leq x\} \ \in \ \mathcal{A}$$

for all $x \in \mathbb{R}$.

Note: Discrete random variables are random variables according to the above definition, and the definition of a *cumulative distribution function* (cdf) given in (3.6) generalizes to all random variables:

$$F_X : \mathbb{R} \to [0,1], \qquad F_X(x) \ := \ \Pr(X \leq x)$$

(4.2) Lemma. From Definition (4.1) it follows, that $X^{-1}(S) \in \mathcal{A}$ for all intervals $S \subseteq \mathbb{R}$.

(4.3) Lemma.

(i) If $A_i \in \mathcal{A}$ with $A_1 \subseteq A_2 \subseteq \ldots$, then: $\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) \ = \ \lim_{n\to\infty} \Pr(A_n)$

(ii) If $A_i \in \mathcal{A}$ with $A_1 \supseteq A_2 \supseteq \ldots$, then: $\Pr\left(\bigcap_{i=1}^{\infty} A_i\right) \ = \ \lim_{n\to\infty} \Pr(A_n)$

(4.4) Lemma. If F_X is a cumulative distribution function, then the following holds:

(i) $\Pr(a < X \leq b) \ = \ F_X(b) - F_X(a)$ for all $a, b \in \mathbb{R}$ with $a < b$

(ii) F_X is monotonically increasing.

(iii) $\lim_{x\to-\infty} F_X(x) \ = \ 0\,, \quad \lim_{x\to\infty} F_X(x) \ = \ 1$

(iv) $F_X(x+) \ := \ \lim_{\xi\to x+} F_X(\xi) \ = \ F_X(x)$ for every $x \in \mathbb{R}$

(v) $\Pr(X = x) \ = \ F_X(x) \ - \ F_X(x-)$ for every $x \in \mathbb{R}$

(4.5) Definition. Let X be a random variable with cdf F_X. The *quantile function* of X is defined for all $p \in (0,1)$ by:

$$F_X^{-1}(p) \ := \ \min\{x \mid F_X(x) \geq p\}$$

First quartile, *median* and *third quartile* are defined to be $F_X^{-1}\left(\frac{1}{4}\right)$, $F_X^{-1}\left(\frac{1}{2}\right)$ and $F_X^{-1}\left(\frac{3}{4}\right)$ respectively.

Note: If F_X is continuous and strictly increasing, then restricting the codomain of F_X to $F_X(\mathbb{R}) = (0,1)$ yields a bijective mapping with the quantile function as the inverse mapping. Moreover:

(i) $F_X\big(F_X^{-1}(p)\big) \ \geq \ p$ for all $p \in (0,1)$

(ii) $F_X\big(F_X^{-1}(p)\big) \ = \ p$ for all $p \in (0,1)$ if F_X is continuous

(iii) F_X^{-1} is strictly increasing if F_X is continuous

(iv) $F_X^{-1}\big(F_X(x)\big) \ \leq \ x$ for all $x \in \mathbb{R}$ with $F_X(x) \in (0,1)$

(4.6) Definition. Any random variable X can be approximated by discrete random variables defined by quantizing the possible values of X:

Any countable set of real numbers

$$\mathcal{P} \;=\; \{x_i \mid i \in \mathbb{Z}\} \qquad \text{with } \; x_{i-1} < x_i \; \text{ for all } \; i \in \mathbb{Z} \; \text{ and } \; \lim_{i\to\pm\infty} x_i = \pm\infty$$

defines a partition of the real line

$$\mathbb{R} \;=\; \dot{\bigcup_{i\in\mathbb{Z}}} (x_{i-1}, x_i]$$

and the *quantizer*

$$q_{\mathcal{P}} : \mathbb{R} \to \mathcal{P}, \qquad q_{\mathcal{P}}(x) \;=\; \min\{x_i \mid x_i \geq x\}$$

Note: If $q_{\mathcal{P}}(x) = x_i$ then:

$$x_{i-1} \;<\; x \;\leq\; x_i \;=\; q_{\mathcal{P}}(x) \qquad \text{and} \qquad 0 \;\leq\; q_{\mathcal{P}}(x) - x \;<\; x_i - x_{i-1}$$

Given any random variable X and $\mathcal{P}$ as above, we may define the discrete random variable $X_{\mathcal{P}} := q_{\mathcal{P}}(X)$ approximating X with values in $\mathcal{P}$ and:

$$p_i \;=\; \Pr(X_{\mathcal{P}} = x_i) \;=\; \Pr(x_{i-1} < X \leq x_i)$$

(4.7) Definition. A random variable X is called a *continuous random variable* if there exists a function $f_X : \mathbb{R} \to \mathbb{R}_0^+$ such that

$$F_X(x) \;=\; \int_{-\infty}^{x} f_X(\xi)\, d\xi$$

for all $x \in \mathbb{R}$. The function f_X is called a *probability density function* (pdf) of X.

(4.8) Lemma. If X is a continuous random variable with pdf f_X, then the following holds:

(i) $\displaystyle\int_{-\infty}^{\infty} f_X(x)\, dx \;=\; 1$

(ii) $\Pr(X = x) \;=\; 0 \quad$ for all $x \in \mathbb{R}$

(iii) $\displaystyle\int_{a}^{b} f_X(x)\, dx \;=\; \Pr(a \leq X \leq b) \quad$ for all $a, b \in \mathbb{R}$, $a < b$

(iv) If f_X is continuous, then:

$$F_X{}'(x) \;=\; f_X(x) \qquad \text{for all } \; x \in \mathbb{R}$$

(4.9) Definition. Let X be a continuous random variable with pdf $f_X(x)$. The *expectation* (*mean*) of X is defined to be

$$E(X) \;:=\; \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx$$

if the integral exists.

(4.10) Remark. Let $\varepsilon > 0$ and

$$\mathcal{P} \;=\; \{x_i \mid i \in \mathbb{Z}\}$$

as in Remark (4.6) with:

$$\Delta(\mathcal{P}) \;:=\; \sup\{x_i - x_{i-1} \mid i \in \mathbb{Z}\} \;\leq\; \varepsilon \qquad (*)$$

Then

$$\begin{aligned}
E(X_{\mathcal{P}}) &= \sum_{i\in\mathbb{Z}} x_i \cdot \Pr(x_{i-1} < X \leq x_i) \\
&= \sum_{i\in\mathbb{Z}} x_i \cdot \int_{x_{i-1}}^{x_i} f_X(x)\,dx \\
&= \sum_{i\in\mathbb{Z}} \int_{x_{i-1}}^{x_i} x_i \cdot f_X(x)\,dx \\
&= \sum_{i\in\mathbb{Z}} \int_{x_{i-1}}^{x_i} q_{\mathcal{P}}(x) \cdot f_X(x)\,dx \\
&= \int_{-\infty}^{\infty} q_{\mathcal{P}}(x) \cdot f_X(x)\,dx \\
&= \int_{-\infty}^{\infty} (x + (q_{\mathcal{P}}(x) - x)) \cdot f_X(x)\,dx \\
&= \int_{-\infty}^{\infty} x \cdot f_X(x)\,dx \;+\; \int_{-\infty}^{\infty} (q_{\mathcal{P}}(x) - x) \cdot f_X(x)\,dx \\
&= E(X) \;+\; \int_{-\infty}^{\infty} (q_{\mathcal{P}}(x) - x) \cdot f_X(x)\,dx
\end{aligned}$$

and:

$$\begin{aligned}
|E(X_{\mathcal{P}}) - E(X)| &= \int_{-\infty}^{\infty} (q_{\mathcal{P}}(x) - x) \cdot f_X(x)\,dx \\
&\leq \int_{-\infty}^{\infty} \varepsilon \cdot f_X(x)\,dx \;=\; \varepsilon \cdot \int_{-\infty}^{\infty} f_X(x)\,dx \;=\; \varepsilon
\end{aligned}$$

Therefore, if $\mathcal{P}_n$ $(n \in \mathbb{N})$ are as in Remark (4.6) with

$$\lim_{n\to\infty} \Delta(\mathcal{P}_n) \;=\; 0$$

then:

$$\lim_{n\to\infty} E(X_{\mathcal{P}_n}) \;=\; E(X)$$

Suitable sets $\mathcal{P}_n$ are for example

$$\mathcal{P}_n \;=\; \left\{ \frac{i}{n} \;\middle|\; i \in \mathbb{Z} \right\} \qquad (n \in \mathbb{N})$$

or:

$$\mathcal{P}_n \;=\; \left\{ \frac{i}{2^n} \;\middle|\; i \in \mathbb{Z} \right\} \qquad (n \in \mathbb{N}_0)$$

(4.11) Remark. Virtually all definitions and basic results from section 3 based on expectations of discrete random variables can be transferred to corresponding definitions and results for continuous random variables.

This can be achieved using discrete approximations $q_{\mathcal{P}_n}(X)$ for a given continuous random variable X, where the quantizers are defined by sets $\mathcal{P}_n$ ($n \in \mathbb{N}$) with $\lim_{n\to\infty} \Delta(\mathcal{P}_n) = 0$.

(4.12) Lemma. Let X be a continuous random variable with pdf $f_X(x)$.

(i) If $g : \mathbb{R} \to \mathbb{R}$ is piecewise continous and $Y = g(X) := g \circ X$, then:

$$E(Y) = \int_{-\infty}^{\infty} g(x) \cdot f_X(x)\, dx$$

(ii) For any $a, b \in \mathbb{R}$:

$$E(a\, X + b) = a\, E(X) + b$$

(4.13) Definition. Let X be a random variable. Then the following numbers and functions are defined:

(i) *n-th moment* of X:

$$E(X^n)$$

(ii) *variance* of X:

$$\sigma^2 := \operatorname{Var}(X) := E((X - E(X))^2) = E(X^2) - (E(X))^2$$

(iii) *standard deviation* of X:

$$\sigma = \sqrt{\operatorname{Var}(X)}$$

(iv) *moment generating function* of X:

$$\phi_X(t) = E\left(e^{tX}\right) = \int_{-\infty}^{\infty} e^{tx} \cdot f_X(x)\, dx \qquad (t \in \mathbb{R})$$

(v) *complex version of moment generating function* of X:

$$\phi_X(s) = E\left(e^{sX}\right) = \int_{-\infty}^{\infty} e^{sx} \cdot f_X(x)\, dx \qquad (s \in \mathbb{C})$$

(vi) *characteristic function* of X:

$$\varphi_X(v) = \phi_X(j \cdot v) = E\left(e^{(j\cdot v)X}\right) \qquad (v \in \mathbb{R})$$

(4.14) Theorem. If there is some $r > 0$, such that $\phi_X(t)$ is defined for all $t \in [-r, r]$, then $\phi_X(s)$ is defined for all complex s with $|s| < r$, all moments $E(X^i)$ are defined and:

$$\phi_X(s) = \sum_{i=0}^{\infty} \frac{E(X^i)}{i!}\, s^i \qquad \text{for } |s| < r$$

(4.15) Remark. Let X be a discrete random variable defined with respect to a pmf $p_X(x)$, where $\{x \mid x \in \mathbb{R}, p_X(x) \neq 0\} = \{x_i \mid i \in I\}$ is countable. Using the δ-function $\delta(x)$, the cdf of X can be expressed as

$$F_X(x) = \int_{-\infty}^{x} \left(\sum_{i \in I} p_X(x_i) \cdot \delta(t - x_i) \right) dt$$

and

$$f_X(x) = \sum_{i \in I} p_X(x_i) \cdot \delta(x - x_i)$$

is a pdf of X.

Transforms of PDF's

(4.16) Remark. $\phi_X(s)$ is the Laplace transform of f_X evaluated at $-s$:

$$\begin{aligned} \phi_X(s) = E(e^{sX}) &= \int_{-\infty}^{\infty} e^{sx} \cdot f_X(x)\,dx \\ &= \int_{-\infty}^{\infty} e^{-(-s)x} \cdot f_X(x)\,dx = (\mathcal{L}(f_X))\,(-s) \end{aligned}$$

(4.17) Remark. $\varphi_X(v)$ is the Fourier transform of f_X evaluated at $-\frac{v}{2\pi}$:

$$\begin{aligned} \varphi_X(v) = E(e^{(jv)X}) &= \int_{-\infty}^{\infty} e^{jvx} \cdot f_X(x)\,dx \\ &= \int_{-\infty}^{\infty} e^{-j \cdot 2\pi(-\frac{v}{2\pi})x} \cdot f_X(x)\,dx = (\mathcal{F}(f_X))\left(-\frac{v}{2\pi}\right) \end{aligned}$$

(4.18) Theorem. The pdf of a random variable X is determined by its characteristic function:

$$f_X(x) = \frac{1}{2\pi} \cdot \left(\mathcal{F}^{-1}(\varphi_X)\right)\left(-\frac{x}{2\pi}\right)$$

(4.19) Theorem. If $(X_i)_{i \in \mathbb{N}}$ is a sequence of random variables, such that the sequence of characteristic functions $(\varphi_{X_i})_{i \in \mathbb{N}}$ converges to some characteristic function φ_X for some random variable X, then $(X_i)_{i \in \mathbb{N}}$ converges to X.

Joint Random Variables

(4.20) Definition. Let X, Y be random variables with respect to the same probability measure Pr, i.e. with respect to the same triple $(\Omega, \mathcal{A}, \Pr)$. The *joint cumulative distribution function* of X and Y, $F_{XY} : \mathbb{R}^2 \to [0,1]$, is defined by:

$$\begin{aligned} F_{XY}(x,y) &:= \Pr(X \le x, Y \le y) \\ &= \Pr(X^{-1}((-\infty, x]) \cap Y^{-1}((-\infty, y])) \qquad \text{for all } x, y \in \mathbb{R} \end{aligned}$$

(4.21) Lemma. Let X, Y be random variables with respect to the same probability measure Pr and $(x_1, x_2] \times (y_1, y_2]$ be a rectangle. Then:

$$\begin{aligned} &\Pr(x_1 < X \le x_2,\ y_1 < Y \le y_2) \\ &\quad = F_{XY}(x_2, y_2) - F_{XY}(x_1, y_2) - F_{XY}(x_2, y_1) + F_{XY}(x_1, y_1) \end{aligned}$$

(4.22) Lemma. Let F_{XY} be the joint cumulative distribution function of two random variables X and Y. The (marginal) cumulative distribution functions of X and Y are determined by F_{XY} as follows:

$$\begin{aligned} F_X(x) &= \lim_{y \to \infty} F_{XY}(x,y) \\ F_Y(y) &= \lim_{x \to \infty} F_{XY}(x,y) \end{aligned}$$

(4.23) Definition. Two random variables X and Y, defined with respect to the same probability measure Pr, are called *independent*, if for all intervals $S, T \subseteq \mathbb{R}$:

$$\Pr(X \in S,\ Y \in T) = \Pr(X \in S) \cdot \Pr(Y \in T)$$

(4.24) Lemma. Let X, Y be random variables with respect to the same probability measure Pr. Then X and Y are independent, if and only if one of the following holds:

(i) For all $x_1, x_2, y_1, y_2 \in \mathbb{R}$ with $x_1 < x_2$ and $y_1 < y_2$:

$$\Pr(x_1 < X \le x_2,\ y_1 < Y \le y_2) = \big(F_X(x_2) - F_X(x_1)\big) \cdot \big(F_Y(y_2) - F_Y(y_1)\big)$$

(ii) For all $x, y \in \mathbb{R}$:

$$F_{XY}(x,y) = F_X(x) \cdot F_Y(y)$$

(4.25) Lemma. Let X, Y be joint independent random variables and $h, k : \mathbb{R} \to \mathbb{R}$ be piecewise continuous functions, then $h(X)$ and $k(Y)$ are independent.

(4.26) Definition. Two random variables X and Y are *jointly continuous random variables*, if there exists a function $f_{XY} : \mathbb{R}^2 \to \mathbb{R}_0^+$, such that

$$F_{XY}(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{XY}(s,t)\, ds\, dt$$

for all $x, y \in \mathbb{R}$. The function f_{XY} is called a *joint probability density function* of X and Y.

(4.27) Lemma. Let f_{XY} be a joint probability density function of two random variables X and Y. The (marginal) probability density functions of X and Y are determined by f_{XY} as follows:

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{XY}(x,t)\, dt \\ f_Y(y) &= \int_{-\infty}^{\infty} f_{XY}(s,y)\, ds \end{aligned}$$

(4.28) Lemma. Let f_{XY} be a joint probability density function of two random variables X and Y. If f_{XY} is continuous, then for all $x, y \in \mathbb{R}$:

$$\frac{\partial^2}{\partial y \partial x} F_{XY}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y) = f_{XY}(x,y)$$

(4.29) Lemma. Two joint continuous random variables X and Y with probability density functions f_X and f_Y are independent, if and only if

$$f(x,y) = f_X(x) \cdot f_Y(y) \qquad \text{for all } x, y \in \mathbb{R}$$

defines a joint density function of X and Y.

(4.30) Lemma. Let f_{XY} be a joint probability density function of two random variables X and Y.

(i) For any (essential) continuous mapping $g : \mathbb{R}^2 \to \mathbb{R}$ the following holds:

$$E(g(X,Y)) \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, f_{XY}(x,y)\, dx\, dy$$

(ii)

$$E(X+Y) \;=\; E(X) + E(Y)$$

(iii) If X and Y are independent and if $h, k : \mathbb{R} \to \mathbb{R}$ are piecewise continuous functions, then:

$$E(h(X)\cdot k(Y)) \;=\; E(h(X))\cdot E(k(Y))$$

(4.31) Remark. The definitions and results from section 3 concerning a pair of random variables ((3.29)–(3.40)) hold in full generality or their proofs make use of the properties given in (4.25), (4.30)(ii) and (iii). Therefore, all of these definitions and results also apply to continuous random variables.

In particular, if

$$X \;=\; X_1 \,+\, X_2 \,+\, \ldots \,+\, X_n$$

then:

(i)

$$E(X) \;=\; E(X_1) + E(X_2) + \ldots + E(X_n)$$

(ii) If $X_1, \ldots, X_n$ are independent, then:

$$\mathrm{Var}(X) \;=\; \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \ldots + \mathrm{Var}(X_n)$$

(iii) If $X_1, \ldots, X_n$ are independent, then:

$$\varphi_X(v) \;=\; \varphi_{X_1}(v)\cdot\varphi_{X_2}(v)\cdots\varphi_{X_n}(v)$$

(iv) If $X_1, \ldots, X_n$ are independent and its moment generating functions exist, then:

$$\phi_X(t) \;=\; \phi_{X_1}(t)\cdot\phi_{X_2}(t)\cdots\phi_{X_n}(t)$$

(4.32) Theorem. If X and Y are independent random variables with pdf's f_X and f_Y, respectively, then the convolution of f_X and f_Y is a pdf of $X+Y$:

$$f_{X+Y}(z) \;=\; (f_X * f_Y)(z) \;=\; \int_{-\infty}^{\infty} f_X(x)\cdot f_Y(z-x)\, dx$$

(4.33) Theorem. Let X and Y be independent continuous random variables with pdf's f_X and f_Y, respectively. A pdf of $X\cdot Y$ is given by:

$$f_{X\cdot Y}(z) \;=\; \int_{-\infty}^{\infty} f_X(x)\cdot f_Y(z/x)\cdot\left|\frac{1}{x}\right| dx \;=\; \int_{-\infty}^{\infty} f_X(z/y)\cdot f_Y(y)\cdot\left|\frac{1}{y}\right| dy$$

(4.34) Theorem. Let X and Y be independent continuous random variables with pdf's f_X and f_Y, respectively. A pdf of $\frac{X}{Y}$ is given by:

$$f_{\frac{X}{Y}}(z) \;=\; \int_{-\infty}^{\infty} f_X(zy)\, f_Y(y)\, |y|\, dy$$

4.1 Uniform distribution

(4.35) Definition. A random variable X with a pdf given by

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

is said to have a *uniform distribution*, $X \sim \text{uniform[a, b]}$.

(4.36) Lemma. If $X \sim \text{uniform[a, b]}$, then:

(i) $$F_X(x) = \begin{cases} 0 & \text{for } x \le a \\ \frac{x-a}{b-a} & \text{for } a < x \le b \\ 1 & \text{for } x > b \end{cases}$$

(ii) $$E(X) = \frac{a+b}{2}, \quad E(X^n) = \frac{b^{n+1} - a^{n+1}}{(n+1)(b-a)} = \frac{\sum_{i=0}^{n} a^{n-i} b^i}{n+1} \quad \text{for all } n \in \mathbb{N}$$

(iii) $$\text{Var}(X) = \frac{(b-a)^2}{12}, \quad \text{skew}(X) = 0, \quad \text{kurt}(X) = \frac{9}{5}$$

(vi) $$\phi_X(t) = \frac{e^{bt} - e^{at}}{(b-a)t}$$

(4.37) Theorem (Probability Integral Transformation). Let X be a continuous random variable with cdf F_X.

(i) The so-called probability integral transformation

$$Y = F_X(X)$$

has a uniform distribution over $[0, 1]$, i.e.:

$$Y \sim \text{uniform}[0, 1]$$

(ii) let U be a uniformly distributed random variable $U \sim \text{uniform}[0, 1]$. Then:

$$F_X = F_{F_X^{-1}(U)}$$

(4.38) Remark. If a random number generator `rand` is available, that generates uniformly distributed numbers from the interval $(0, 1)$, then (4.37)(ii) may be applied to construct a random number generator with a distribution given by any cdf F_X: Simply apply F_X^{-1} to the output sequence of `rand`.

4.2 Exponential distribution

(4.39) Definition. A random variable X has an *exponential distribution*, $X \sim \exp(\lambda)$, for some *rate parameter* $\lambda > 0$, if it has a pdf defined by:

$$f_X(x) \;=\; \begin{cases} \lambda e^{-\lambda x} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$$

$$\left(\text{Note:} \quad \int_0^\infty \lambda e^{-\lambda x}\, dx \;=\; -e^{-\lambda x}\Big|_0^\infty \;=\; 1\right)$$

(4.40) Lemma. If $X \sim \exp(\lambda)$, then:

(i) $F_X(x) \;=\; \begin{cases} 1 - e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$

(ii) $E(X^n) \;=\; \dfrac{n!}{\lambda^n}$ for $n \in \mathbb{N}$

(iii) $E(X) \;=\; \dfrac{1}{\lambda}$, $\operatorname{Var}(X) \;=\; \dfrac{1}{\lambda^2}$, $\operatorname{skew}(X) \;=\; 2$, $\operatorname{kurt}(X) \;=\; 9$

(iv) $\phi_X(t) \;=\; \dfrac{\lambda}{\lambda - t}$ for $t < \lambda$

(v) $F_X^{-1}(p) \;=\; -\dfrac{\ln(1-p)}{\lambda}$ for $p \in (0,1)$

(vi) first quartile, median and third quartile are:

$\ln(4/3)/\lambda \approx 0.288/\lambda$, $\ln(2)/\lambda \approx 0.693/\lambda$, $\ln(4)/\lambda \approx 1.386/\lambda$

(4.41) Theorem. X has an exponential distribution, if and only if X has the memoryless property:

$$\Pr(X > t + h \mid X > t) \;=\; \Pr(X > h) \qquad \text{for all } t, h \geq 0$$

(4.42) Theorem. Let T be a continous random variable with a continous pdf f_T with constant success rate λ, i. e. for all $t > 0$, we have:

$$\lambda \;=\; \lim_{\Delta t \to 0} \frac{\Pr(T \leq t + \Delta t \mid T > t)}{\Delta t}$$

Then $T \sim \exp(\lambda)$.

(4.43) Theorem. Let $X_1, \ldots, X_n$ be independent random variables with $X_i \sim \exp(\lambda_i)$ for $i = 1, \ldots, n$. Then

$$X \;=\; \min\{X_1, \ldots, X_n\}$$

is a random variable with $X \sim \exp(\lambda_1 + \cdots + \lambda_n)$.

(4.44) Theorem. Let $\lambda > 0$ and let $X_1, \ldots, X_n$ be independent random variables with $X_i \sim \exp(\lambda)$ for $i = 1, \ldots, n$. Furthermore, let $X_{i_1}, \ldots, X_{i_n}$ be a reordering of $X_1, \ldots, X_n$, such that:

$$X_{i_1} \;\leq\; \ldots \;\leq\; X_{i_n}$$

Then

$$Y_k \;:=\; X_{i_k} - X_{i_{k-1}}$$

is random variable with $Y_k \sim \exp((n+1-k)\,\lambda)$.

Reliability theory

(4.45) Definition. Let T be a random variable describing the *lifetime* or *time-to-failure* of a device or system. Let f_T and F_T be the pdf and cdf of T, respectively.

(i) The *reliability function* is defined by:

$$R(t) \;:=\; \Pr(T > t) \;=\; 1 \;-\; F_T(t)$$

(ii) The *mean time to failure* (*MTTF*) is defined to be the expected lifetime:

$$\text{MTTF} \;:=\; E(T)$$

(iii) The *failure rate* is defined to be

$$\lambda(t) \;:=\; \lim_{\Delta t \to 0} \frac{\Pr(T \le t + \Delta t \mid T > t)}{\Delta t}$$

Since lifetimes are non-negative random variables, the MTTF equals the integral of the reliability function:

(4.46) Lemma.

$$\text{MTTF} \;=\; \int_0^\infty R(t)\,dt$$

Proof:

$$\begin{aligned}
\text{MTTF} \;=\; E(T) \;&=\; \int_0^\infty t \cdot f_T(t)\,dt \\
&=\; -t \cdot \big(1 - F_T(t)\big)\Big|_0^\infty \;+\; \int_0^\infty \big(1 - F_T(t)\big)\,dt \\
&=\; \int_0^\infty \big(1 - F_T(t)\big)\,dt \;=\; \int_0^\infty R(t)\,dt
\end{aligned}$$

□

(4.47) Lemma.

(i)

$$\lambda(t) \;=\; \frac{f_T(t)}{1 - F_T(t)} \;=\; \frac{f_T(t)}{\displaystyle\int_t^\infty f_T(\tau)\,d\tau} \;=\; -\frac{R'(t)}{R(t)}$$

(ii)

$$F_T(t) \;=\; 1 - e^{-\int_0^t \lambda(\tau)\,d\tau}, \qquad f_T(t) \;=\; \lambda(t)\,e^{-\int_0^t \lambda(\tau)\,d\tau}$$

(iii) If $\lambda(t) = \lambda$ is constant, then $T \sim \exp(\lambda)$.

Proof:

(i)

$$\begin{aligned}
\lim_{\Delta t \to 0} \frac{\Pr(T \leq t + \Delta t \mid T > t)}{\Delta t} &= \lim_{\Delta t \to 0} \frac{\frac{\Pr(t < T \leq t+\Delta t)}{\Pr(T > t)}}{\Delta t} \\
&= \frac{1}{\Pr(T > t)} \lim_{\Delta t \to 0} \frac{\Pr(t < T \leq t + \Delta t)}{\Delta t} \\
&= \frac{1}{1 - F_T(t)} \lim_{\Delta t \to 0} \frac{F_T(t + \Delta t) - F_T(t)}{\Delta t} \\
&= \frac{1}{1 - F_T(t)} \cdot F_T'(t) \\
&= \frac{f_T(t)}{1 - F_T(t)}
\end{aligned}$$

(ii)

$$\begin{aligned}
\lambda(t) = \frac{F_T'(t)}{1 - F_T(t)} \quad &\Longrightarrow \quad \int_0^t \lambda(\tau)\, d\tau = \int_0^t \frac{F_T'(\tau)}{1 - F_T(\tau)}\, d\tau = -\ln(1 - F_T(\tau))\Big|_0^t \\
&\Longrightarrow \quad \int_0^t \lambda(\tau)\, d\tau = -\ln(1 - F_T(t)) \\
&\Longrightarrow \quad e^{-\int_0^t \lambda(\tau)\, d\tau} = 1 - F_T(t) \\
&\Longrightarrow \quad F_T(t) = 1 - e^{-\int_0^t \lambda(\tau)\, d\tau} \\
&\Longrightarrow \quad f_T(t) = \lambda(t)\, e^{-\int_0^t \lambda(\tau)\, d\tau}
\end{aligned}$$

(iii) Follows from (ii).

□

4.3 Normal distribution

(4.48) Theorem.

(i) $$\int_{-\infty}^{\infty} e^{-x^2}\,dx \;=\; \sqrt{\pi}$$

(ii) $$\frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\,dx \;=\; 1 \qquad \text{for all } \mu,\sigma\in\mathbb{R} \text{ with } \sigma>0$$

(4.49) Definition. A random variable has a *normal distribution* $\mathcal{N}(\mu,\sigma)$ for some parameters $\mu,\sigma\in\mathbb{R}$, $\sigma>0$ if it has a pdf defined by:

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

(4.50) Definition. The normal distribution $\mathcal{N}(0,1)$ is called the *standard normal distribution.* Its pdf and cdf are denoted by

$$\varphi(x) \;:=\; \frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{1}{2}x^2}$$

and

$$\Phi(x) \;:=\; \frac{1}{\sqrt{2\pi}}\cdot\int_{-\infty}^{x} e^{-\frac{1}{2}u^2}\,du$$

respectively. Graphs of $\varphi(x)$ and $\Phi(x)$ are shown in Figure 19.

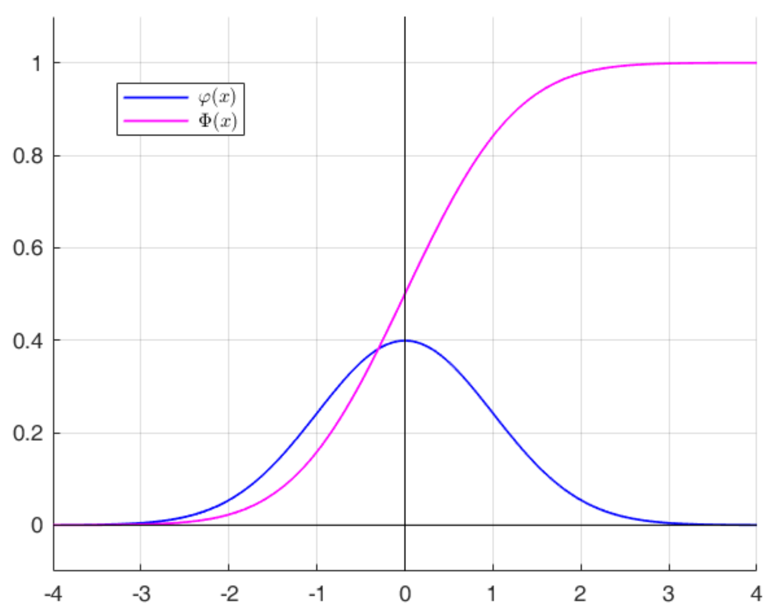

Figure 19: Pdf $\varphi(x)$ and cdf $\Phi(x)$ of the standard normal distribution

(4.51) Lemma. If $X \sim \mathcal{N}(\mu, \sigma)$, then:

(i) $E(X) \;=\; \mu$

(ii) $\operatorname{Var}(X) \;=\; \sigma^2$

(iii) $\phi_X(t) \;=\; e^{\mu t + \frac{1}{2}\sigma^2 t^2}$ (for all $t \in \mathbb{R}$)

(iv) $aX + b \;\sim\; \mathcal{N}(a\mu + b,\, |a| \cdot \sigma)$ if $a \neq 0$

(v) $F_X(x) \;=\; \Phi\left(\dfrac{x - \mu}{\sigma}\right)$

(vi) $\Phi(-x) \;=\; 1 - \Phi(x)$ for all $x \in \mathbb{R}$

(vii) $\Phi^{-1}(p) \;=\; -\Phi^{-1}(1-p)$ for all $p \in (0, 1)$

(4.52) MatLab. MatLab provides implementations of pdf's, cdf's and the inverses of cdf's of normal distributions under the names `normpdf()`, `normcdf()` and `norminv()`, respectively.

(4.53) Theorem. If $X_1, \ldots, X_n$ are independent random variables with

$$X_i \sim \mathcal{N}(\mu_i, \sigma_i)$$

then

$$X \;=\; X_1 + \cdots + X_n$$

has a normal distribution with mean $\mu = \mu_1 + \cdots + \mu_n$ and variance $\sigma^2 = {\sigma_1}^2 + \cdots + {\sigma_n}^2$.

(4.54) Theorem (Central limit theorem). Let $(X_i)_{i \in \mathbb{N}}$ be a sequence of independent, identically distributed random variables with finite mean μ and finite variance σ^2. Let F_{Y_n} be the cdf of

$$Y_n \;=\; \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right) \;=\; \frac{\left(\sum_{i=1}^{n} X_i\right)/n \;-\; \mu}{\sigma/\sqrt{n}}$$

Then $E(Y_n) = 0$, $\operatorname{Var}(Y_n) = 1$ for all $n \in \mathbb{N}$ and:

$$\lim_{n \to \infty} F_{Y_n} \;=\; \Phi$$

4.4 Gamma distribution

(4.55) Definition. For $\alpha > 0$ the *gamma function* is defined by:

$$\Gamma(\alpha) \;=\; \int_0^\infty x^{\alpha-1} e^{-x}\, dx$$

(4.56) Theorem.

(i) $\Gamma(\alpha) \;=\; (\alpha-1)\,\Gamma(\alpha-1) \qquad \text{for all } \alpha > 1$

(ii) $\Gamma(n) \;=\; (n-1)! \qquad \text{for all } n \in \mathbb{N}$

(iii) $\Gamma(\tfrac{1}{2}) \;=\; \sqrt{\pi}$

(iv) $\displaystyle\int_0^\infty x^{\alpha-1}\, e^{-\beta x}\, dx \;=\; \frac{\Gamma(\alpha)}{\beta^\alpha} \qquad \text{for all } \alpha, \beta > 0$

(4.57) Definition. A random variable has a *gamma distribution* $\Gamma(\alpha, \beta)$ for some parameters $\alpha, \beta \in \mathbb{R}^+$ if its pdf is defined by:

$$f_X(x) \;=\; \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-\beta x} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$$

(4.58) Theorem. If X has a gamma distribution with $X \sim \Gamma(\alpha, \beta)$, then:

(i) $E(X) \;=\; \dfrac{\alpha}{\beta}$

(ii) $\mathrm{Var}(X) \;=\; \dfrac{\alpha}{\beta^2}$

(iii) $\phi_X(t) \;=\; \left(\dfrac{\beta}{\beta - t}\right)^\alpha \qquad \text{for } t < \beta$

(iv) Scaling property of Gamma distributions:

$b \cdot X \;\sim\; \Gamma(\alpha, \beta/b) \qquad \text{for } b \in \mathbb{R}^+$

(4.59) MatLab. MatLab provides implementations of pdf's, cdf's and the inverses of cdf's of gamma distributions under the names `gampdf()`, `gamcdf()` and `gaminv()`, respectively.

(4.60) Theorem. If $X_1, \ldots, X_n$ are independent random variables with $X_i \sim \Gamma(\alpha_i, \beta)$, then

$$X \;=\; X_1 + \cdots + X_n$$

has a gamma distribution with $X \sim \Gamma\left(\displaystyle\sum_{i=1}^n \alpha_i,\ \beta\right)$.

4.5 Erlang distributions

(4.61) Definition. A random variable has an *Erlang distribution* for some parameters $n \in \mathbb{N}$ and $\lambda \in \mathbb{R}^+$ if its pdf is defined by:

$$f_X(x) \;=\; \begin{cases} \frac{\lambda^n}{(n-1)!}\, x^{n-1}\, e^{-\lambda x} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$$

I.e.:

$$\text{Erlang}\,(n,\lambda) \;=\; \Gamma(n,\lambda)$$

(4.62) Remark.

$$\exp(\lambda) \;=\; \text{Erlang}\,(1,\lambda) \;=\; \Gamma(1,\lambda)$$

(4.63) Theorem. If X has an Erlang distribution with $X \sim$ Erlang (n,λ), then:

(i) $E(X) \;=\; \dfrac{n}{\lambda}$

(ii) $\text{Var}(X) \;=\; \dfrac{n}{\lambda^2}$

(iii) $\phi_X(t) \;=\; \left(\dfrac{\lambda}{\lambda - t}\right)^n \qquad \text{for } t < \lambda$

(iv) Scaling property of Erlang distributions:

$b \cdot X \;\sim\; \text{Erlang}\,(n, \lambda/b) \qquad \text{for } b \in \mathbb{R}^+$

In particular: $2\lambda \cdot X \sim \text{Erlang}\,(n, 1/2) \;=\; \Gamma(n, 1/2) \;=\; \chi^2_{2n}$
(See also Definition (4.73) and Remark (4.74)).

(4.64) Theorem. If $X_1, \ldots, X_n$ are independent random variables with $X_i \sim \exp(\lambda)$, then

$$X \;=\; X_1 + \cdots + X_n$$

has an Erlang distribution with $X \sim$ Erlang (n,λ).

Pdf's of Erlang$(n,1)$ distributions for small values of n are shown in Figure 20.

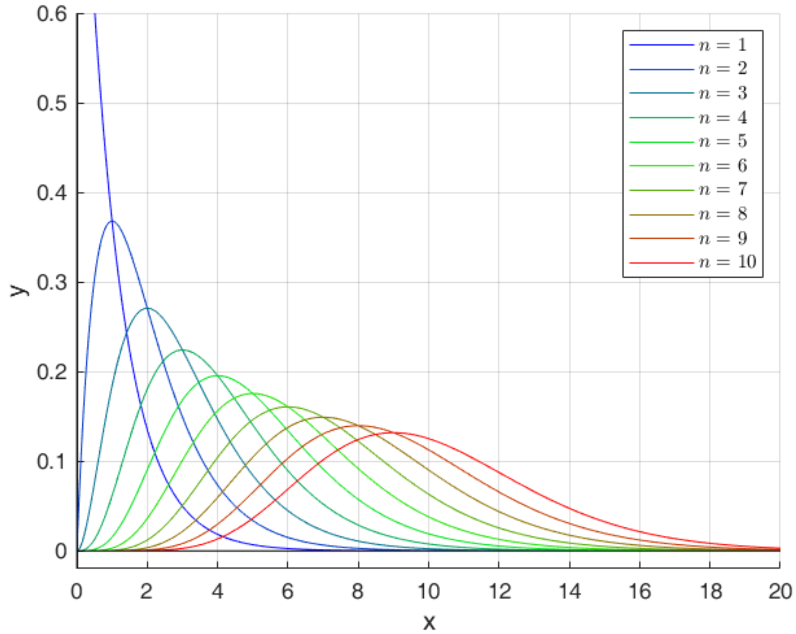

Figure 20: Pdf's of Erlang$(n, 1) = \Gamma(n, 1)$ distributions for $n = 1, 2, \ldots, 10$

(4.65) Lemma. Let f_X be the pdf of a random variable $X \sim$ Erlang$\,(n, \lambda)$. Then

(i) f_X has a maximum at: $m_n \;=\; \dfrac{n-1}{\lambda}$

(ii) If $n = 1$, then f_X is concave upward.

(iii) If $n = 2$, then f_X is concave downward and then upward, with inflection point at: $\dfrac{2}{\lambda}$

(iv) If $n > 2$, then f_X is concave upward, then downward, then upward again with inflection points at: $m_n \pm \dfrac{\sqrt{n-1}}{\lambda}$

(4.66) Theorem. If X has an Erlang distribution with $X \sim$ Erlang$\,(n, \lambda)$, then its cdf is given by:

$$F_X(x) \;=\; \begin{cases} 1 \;-\; \displaystyle\sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!}\, e^{-\lambda x} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$$

Poisson processes

(4.67) Definition. A process in which discrete similar events occur randomly in time, can be described by the following sequences of random variables:

- *Arrival times*: $T = (T_1, T_2, \dots)$ (also set $T_0 := 0$)
- *Inter-arrival times*: $X = (X_1, X_2, \dots)$ with $X_n := T_n - T_{n-1}$

$$\left(\implies \quad T_n \; = \; \sum_{i=1}^{n} X_i \;\text{ for all }\; n \in \mathbb{N} \right)$$

Such a process can also be described with the following set of random variables:

- *Counting process*: $N = \{N_t \mid t \geq 0\}$ with $N_t := \max\{n \in \mathbb{N}_0 \mid T_n \leq t\}$

$$\left(\implies \quad T_n \; = \; \min\{t \geq 0 \mid N_t = n\} \;\text{ for all }\; n \in \mathbb{N}_0 \right)$$

(4.68) Lemma. The following are equivalent:

(i) At least n arrivals occurred in the interval $(0, t]$.

(ii) $N_t \; \geq \; n$

(iii) $T_n \; \leq \; t$

(4.69) Definition.

- A process in which events occur randomly in time is called a *renewal process*, if the inter-arrival times $X_1, X_2, \dots$ are independent, identically distributed random variables.
- A renewal process is called a *Poisson process*, if it satisfies the *strong* renewal assumption, that at each fixed time, the process restarts probabilistically, independent of the past.

(4.70) Theorem. Given a Poisson process with arrival times $T = (T_0 = 0, T_1, T_2, \dots)$, inter-arrival times $X = (X_1, X_2, \dots)$ and counting variables $N = \{N_t \mid t \geq 0\}$, there exists some parameter $\lambda > 0$, such that:

- $X_n \; \sim \; \exp(\lambda)$ for all $n \in \mathbb{N}$
- $T_n \; \sim \; \text{Erlang}\,(n, \lambda)$ for all $n \in \mathbb{N}$
- $N_t \; \sim \; \text{Poisson}(\lambda t)$ for all $t > 0$

Proof: For a Poisson process the X_n's are i.i.d. random variables and have the *memoryless property* by the strong renewal assumption. Therefore, it follows from Theorem (4.41) that the X_n's are independent random variables with

$$X_n \; \sim \; \exp(\lambda)$$

for some rate parameter $\lambda > 0$. Moreover, it follows from Theorem (4.64), that:

$$T_n \;\; = \;\; \sum_{i=1}^{n} X_i \; \sim \; \text{Erlang}\,(n, \lambda)$$

Finally, by Lemma (4.68) and Theorem (4.66):

$$\Pr(N_t \geq n) \;=\; \Pr(T_n \leq t) \;=\; F_{T_n}(t) \;=\; 1 - \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!}\, e^{-\lambda t}$$

$$\begin{aligned}
\Longrightarrow \qquad \Pr(N_t = n) &= \Pr(N_t \geq n) - \Pr(N_t \geq n+1) \\
&= 1 - \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!}\, e^{-\lambda t} - \left(1 - \sum_{k=0}^{n} \frac{(\lambda t)^k}{k!}\, e^{-\lambda t}\right) \\
&= \frac{(\lambda t)^n}{n!}\, e^{-\lambda t}
\end{aligned}$$

$$\Longrightarrow \qquad N_t \sim \text{Poisson}(\lambda t)$$

□

(4.71) Definition. A Poisson process is said to have *rate* λ, if the inter-arrival times have an exponential distribution with rate parameter λ.

(4.72) Remark. A Bernoulli trials process $(B_t)_{t\in\mathbb{N}}$ with independent random variables $B_t \sim \text{Bernoulli}(p)$ can be considered to be a discrete time version of the Poisson process with:

- *Counting process*: $N = (N_t)_{t\in\mathbb{N}_0}$ with

$$N_t \;=\; \sum_{i=1}^{t} B_i \;\sim\; \text{binomial}(t,p)$$

  $N = (N_t)_{t\in\mathbb{N}_0}$ has independent, stationary increments ($N_{t_2} - N_{t_1} = N_{t_2-t_1}$ for all $t_1, t_2 \in \mathbb{N}_0$, $t_1 \leq t_2$)

- *Inter-arrival times*: $X = (X_1, X_2, \dots)$ with independent random variables

$$X_i \;\sim\; \text{geometric}(p)$$

- *Arrival times*: $T = (T_1, T_2, \dots)$ has independent, stationary increments and

$$T_i \;\sim\; \text{nbino}(i,p)$$

4.6 Chi-Square distributions

(4.73) Definition. A random variable has a *chi-square distribution with n degrees of freedom* χ^2_n if its pdf is defined by

$$f_X(x) \quad = \quad \begin{cases} \frac{1}{2^{\frac{n}{2}}\cdot\Gamma(\frac{n}{2})}\, x^{\frac{n}{2}-1}\, e^{-\frac{x}{2}} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$$

i.e.:

$$\chi^2_n \quad = \quad \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$$

Pdf's and cdf's of χ^2_n distributions for small values of n are shown in Figure 21. Figure 22 shows pdf's of some χ^2_n distributions in comparison with pdf's of normal distributions (dashed lines) having the same expectations and standard deviations. Finally, Figure 23 shows pdf's of some normalized χ^2_n distributions, i. e. for some $n \in \mathbb{N}$ and $X_n \sim \chi^2_n$ the pdf's $f_{X^o_n}$ of $X^o_n := \frac{X-n}{\sqrt{n}}$ are shown. Note, how these converge to the pdf $\varphi(x)$ of the standard normal distribution.

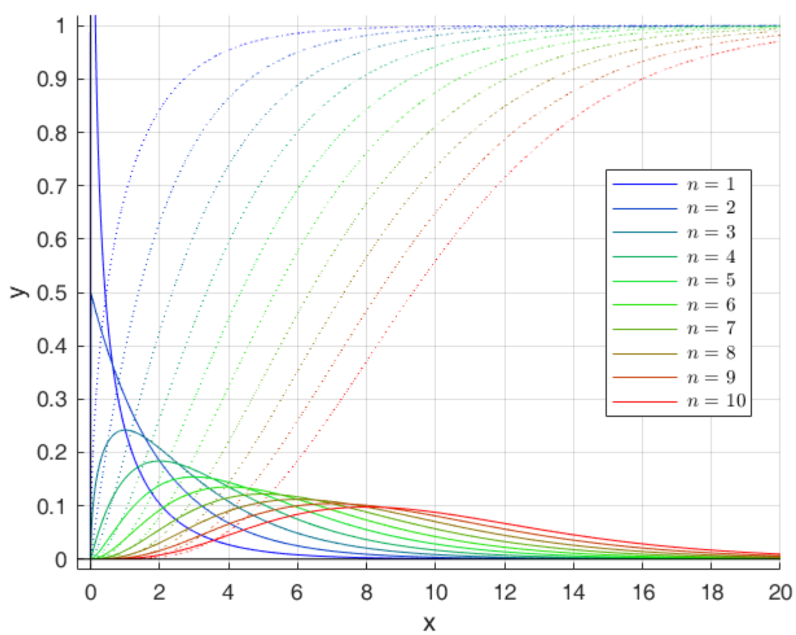

Figure 21: Pdf's and cdf's of $\chi^2_n = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$ distributions for $n = 1, 2, \ldots, 10$

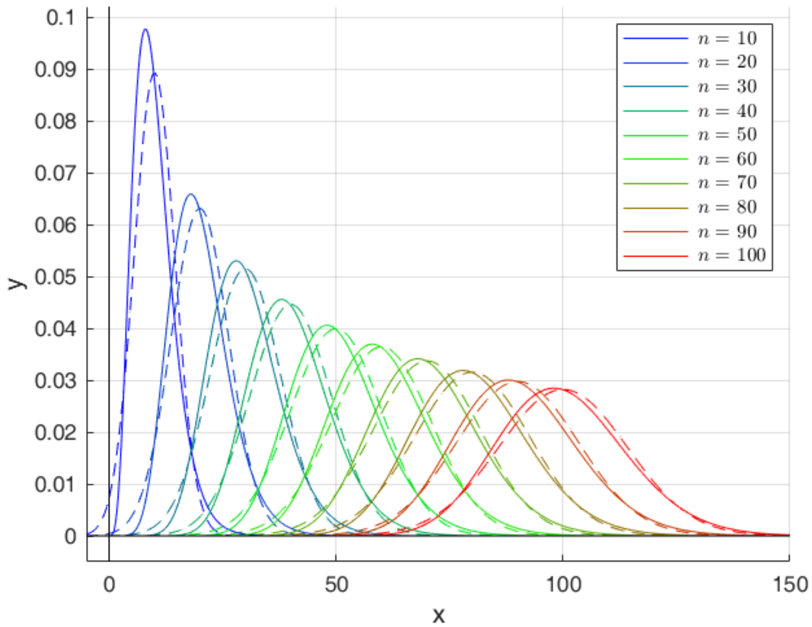

Figure 22: Pdf's of χ^2_n and $\mathcal{N}(n, \sqrt{2n})$ distributions for $n = 10, 20, \ldots, 100$

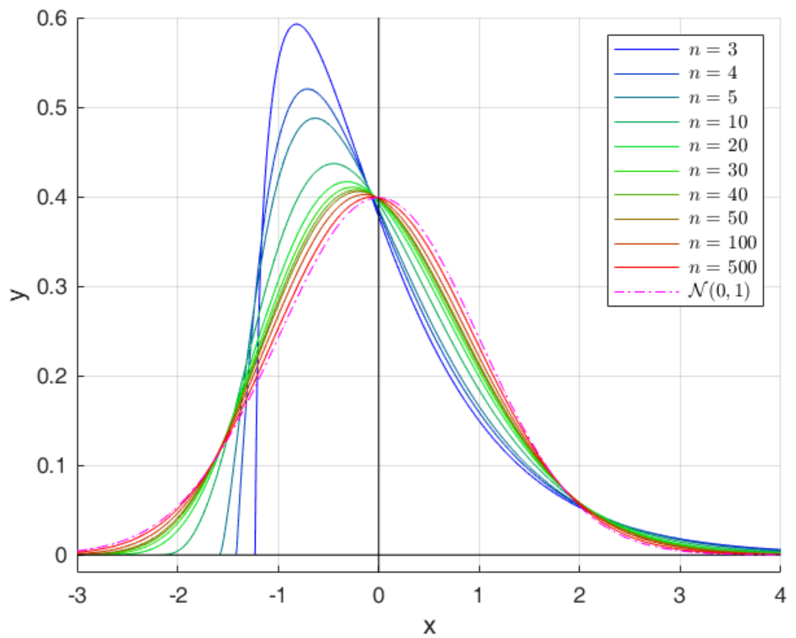

Figure 23: Pdf's of normalized χ^2_n distributions for selected values of n and the pdf of the standard normal distribution

(4.74) Remark. From Theorem 63 (iv) it follows that pdf's of χ^2_{2n} distributions are scaled versions of pdf's of Erlang$(n, 1)$ for all $n \in \mathbb{N}$. (Check this for $n = 1, \ldots, 5$ by comparing the pdf's in Figure 20 with the corresponding pdf's in Figure 21.)

(4.75) Theorem. If X has a chi-square distribution with n degrees of freedom, i.e. $X \sim \chi^2_n$, then:

(i) $E(X) \;=\; n$

(ii) $\mathrm{Var}(X) \;=\; 2n$

(iii) $\phi_X(t) \;=\; \left(\dfrac{1}{1-2t}\right)^{\frac{n}{2}} \qquad \text{for } \; t < 1/2$

(4.76) MatLab. MatLab provides implementations of pdf's, cdf's and the inverses of cdf's of chi-square distributions under the names `chi2pdf()`, `chi2cdf()` and `chi2inv()`, respectively.

(4.77) Theorem. If $X_1, \ldots, X_m$ are independent random variables with $X_i \sim \chi^2_{n_i}$, then

$$X \;=\; X_1 + \cdots + X_m$$

has a chi-square distribution with $X \sim \chi^2_n$, where $n = \sum_{i=1}^m n_i$.

(4.78) Theorem. If $Z_1, \ldots, Z_n$ are independent random variables with $Z_i \sim \mathcal{N}(0, 1)$ for all i, then

$$X \;=\; Z_1^2 + \cdots + Z_n^2$$

has a chi-square distribution with n degrees of freedom, i.e. $X \sim \chi^2_n$.

(4.79) Theorem. If $X_1, \ldots, X_n$ are independent random variables with $X_i \sim \mathcal{N}(\mu, \sigma)$ for all i, then

$$\widehat{\sigma_0{}^2} \;=\; \frac{1}{n}\left((X_1 - \mu)^2 + \cdots + (X_n - \mu)^2\right)$$

has a gamma distribution with $\widehat{\sigma_0{}^2} \sim \Gamma\left(\frac{n}{2}, \frac{n}{2\sigma^2}\right)$.

4.7 t-distributions

(4.80) Definition. If Z and Y_n are independent random variables with $Z \sim \mathcal{N}(0,1)$ and $Y_n \sim \chi^2_n$, then the distribution of

$$T_n \quad := \quad \frac{Z}{\sqrt{Y_n/n}}$$

is called a *t-distribution with n degrees of freedom*, denoted by:

$$T_n \sim t_n$$

The cdf of a t-distribution with n degrees of freedom is denoted by:

$$F_{t_n}$$

(4.81) Theorem. If $T_n \sim t_n$, then a pdf of T_n is given by:

$$f_{t_n}(x) \quad = \quad \frac{\left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}}{\sqrt{n} \cdot B\left(\frac{1}{2}, \frac{n}{2}\right)}$$

(4.82) Remark. The pdf f_{t_n} of a t-distribution is an even function and therefore:

$$F_{t_n}(x) \quad = \quad 1 \; - \; F_{t_n}(-x) \qquad \text{for all} \;\; x \in \mathbb{R}$$

Figure 24 shows pdf's and cdf's of t_n distributions for small values of n and the density functions of the standard normal distribution.

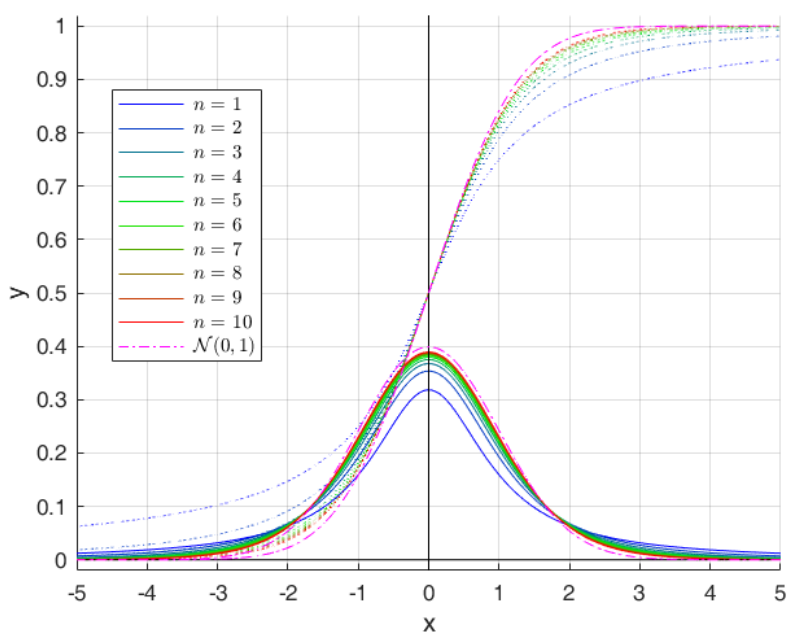

Figure 24: Pdf's and cdf's of t_n distributions for $n = 1, 2, \ldots, 10$ and of the standard normal distribution

(4.83) Theorem. If T_n has a t-distribution $T_n \sim t_n$, then:

(i) $E({T_n}^k)$ exists, if and only $k < n$.

(ii) If $k < n$ is odd, then $E({T_n}^k) = 0$.

(iii) If $k < n$ is even, then:

$$E({T_n}^k) \quad = \quad n^{\frac{k}{2}} \cdot \frac{\Gamma\left(\frac{k+1}{2}\right) \cdot \Gamma\left(\frac{n-k}{2}\right)}{\Gamma\left(\frac{1}{2}\right) \cdot \Gamma\left(\frac{n}{2}\right)}$$

(iv) If $n > 2$, then: $E(T_n) = 0$ and $\mathrm{Var}(T_n) = \frac{n}{n-2}$

(4.84) MatLab. MatLab provides implementations of pdf's, cdf's and the inverses of cdf's of t-distributions under the names `tpdf()`, `tcdf()` and `tinv()`, respectively.

4.8 Beta distribution

(4.85) Definition. For $x, y > 0$ the *beta function* is defined by:

$$B(x,y) \;:=\; \int_0^1 t^{x-1}(1-t)^{y-1}\,dt$$

(4.86) Theorem. For all $x, y > 0$:

$$B(x,y) \;=\; \frac{\Gamma(x)\cdot\Gamma(y)}{\Gamma(x+y)}$$

(4.87) Definition. A random variable has a *beta distribution* $\text{beta}(\alpha_1, \alpha_2)$ for some parameters $\alpha_1, \alpha_2 \in \mathbb{R}^+$, if its pdf is defined by:

$$f_X(x) \;=\; \begin{cases} \frac{1}{B(\alpha_1,\alpha_2)}\, x^{\alpha_1-1}(1-x)^{\alpha_2-1} & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

(4.88) MatLab. MatLab provides implementations of pdf's, cdf's and the inverses of cdf's of beta distributions under the names `betapdf()`, `betacdf()` and `betainv()`, respectively.

(4.89) Lemma.

(i) $B(\alpha_1,\alpha_2) \;=\; \displaystyle\int_0^\infty (1-e^{-\theta})^{\alpha_1-1}e^{-\alpha_2\theta}\,d\theta$

(ii) $B(\alpha_1,\alpha_2) \;=\; \displaystyle\int_0^\infty \frac{z^{\alpha_1-1}}{(1+z)^{\alpha_1+\alpha_2}}\,dz$

(iii) $B\left(\dfrac{1}{2}, \dfrac{n+1}{2}\right) \;=\; 2\cdot\displaystyle\int_0^{\pi/2} \sin^n(\theta)\,d\theta$

(4.90) Examples. If X has a beta distribution $X \sim \text{beta}(\alpha_1, \alpha_2)$, then:

(i) $\alpha_1 \;=\; \alpha_2 \;=\; 1 \quad\Longrightarrow\quad X \sim \text{uniform}([0,1])$

(ii) $\alpha_1 \;=\; \alpha_2 \;=\; 2 \quad\Longrightarrow\quad f_X(x) \;=\; -6(x^2-x)\cdot I_{(0,1)}$

(iii) $\alpha_1 \;=\; \frac{1}{2},\;\; \alpha_2 \;=\; 1 \quad\Longrightarrow\quad f_X(x) \;=\; \frac{1}{2\sqrt{x}}\cdot I_{(0,1)}$

(4.91) Exercise. Use (4.86) to provide a new proof of (4.48)(i), i.e.:

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx \;=\; \sqrt{\pi}$$

(4.92) Exercise. Prove: If X and Y are independent continuous random variables with $X \sim \Gamma(\alpha_1, \beta)$ and $Y \sim \Gamma(\alpha_2, \beta)$, then a pdf of $Z = \frac{X}{Y}$ is given by:

$$f_Z(z) \;=\; \frac{1}{B(\alpha_1,\alpha_2)}\cdot\frac{z^{\alpha_1-1}}{(1+z)^{\alpha_1+\alpha_2}}\cdot I_{(0,\infty)}$$

$X \;\sim$	pdf $\;f_X(x)$	cdf $\;F_X(x)$	$E(X)$	$\mathrm{Var}(X)$	$\phi_X(t)$	$X_1+\cdots+X_m \quad (X_i \sim X)$
uniform[a, b]	$\frac{1}{b-a}\cdot I_{[a,b]}$	$\frac{x-a}{b-a}\cdot I_{[a,b]} \;+\; 1\cdot I_{(b,\infty)}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt}-e^{at}}{(b-a)t}$	
$\exp(\lambda)$ $(\lambda>0)$	$\lambda e^{-\lambda x}\cdot I_{[0,\infty)}$	$(1-e^{-\lambda x})\cdot I_{[0,\infty)}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda-t} \quad (t<\lambda)$	Erlang (m,λ)
$\mathcal{N}(\mu,\sigma)$ $\mathcal{N}(0,1)$	$\frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$ $\varphi(x)=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{1}{2}x^2}$ `normpdf(x)`	$F_X(x)=\Phi\left(\frac{x-\mu}{\sigma}\right)$ $\Phi(x)=\int_{-\infty}^{x}\varphi(u)\,du$ `normcdf(x), norminv(p)`	μ 0	σ^2 1	$e^{\mu t+\frac{1}{2}\sigma^2t^2}$ $e^{\frac{1}{2}t^2}$	$\mathcal{N}(m\mu,\sqrt{m}\,\sigma)$
$\Gamma(\alpha,\beta)$ $(\alpha,\beta\in\mathbb{R}^+)$	$\frac{\beta^\alpha}{\Gamma(\alpha)}\,x^{\alpha-1}\,e^{-\beta x}\cdot I_{(0,\infty)}$ `gampdf(x,alpha,1/beta)`	`gamcdf(x,alpha,1/beta)` `gaminv(p,alpha,1/beta)`	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	$\left(\frac{\beta}{\beta-t}\right)^\alpha \quad (t<\beta)$	$\Gamma(m\alpha,\,\beta)$
Erlang $(n,\beta)=\Gamma(n,\beta)$ $(n\in\mathbb{N},\beta\in\mathbb{R}^+)$	$\frac{\beta^n}{(n-1)!}\,x^{n-1}\,e^{-\beta x}\cdot I_{(0,\infty)}$		$\frac{n}{\beta}$	$\frac{n}{\beta^2}$	$\left(\frac{\beta}{\beta-t}\right)^n \quad (t<\beta)$	Erlang (mn,β)
$\chi^2_n=\Gamma\left(\frac{n}{2},\frac{1}{2}\right)$	$\frac{1}{2^{\frac{n}{2}}\cdot\Gamma(\frac{n}{2})}\,x^{\frac{n}{2}-1}\,e^{-\frac{x}{2}}\cdot I_{(0,\infty)}$	`chi2cdf(x,n)`, `chi2inv(p,n)`	n	$2n$	$\left(\frac{1}{1-2t}\right)^{\frac{n}{2}} \quad (t<\frac{1}{2})$	χ^2_{mn}
t_n	$\frac{1}{\sqrt{n}\cdot B\left(\frac{1}{2},\frac{n}{2}\right)}\cdot\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}$	`tcdf(x,n)`, `tinv(p,n)`	0	$\frac{n}{n-2} \quad (n>2)$		
Laplace(λ)	$\frac{\lambda}{2}\,e^{-\lambda\lvert x\rvert}$	$\frac{e^{\lambda x}}{2}\cdot I_{(-\infty,0]}+\left(1-\frac{e^{-\lambda x}}{2}\right)\cdot I_{(0,\infty)}$	0	$\frac{2}{\lambda^2}$	$\frac{\lambda^2}{\lambda^2-t^2} \quad (\lvert t\rvert<\lambda)$	
Cauchy(λ)	$\frac{\lambda/\pi}{\lambda^2+x^2}$	$\frac{1}{2}+\frac{\arctan(x/\lambda)}{\pi}$	n.d.	n.d.	n.d.	
Rayleigh(λ)	$\frac{x}{\lambda^2}\,e^{\frac{-(x/\lambda)^2}{2}}\cdot I_{[0,\infty)}$	$\left(1-e^{-\frac{(x/\lambda)^2}{2}}\right)\cdot I_{[0,\infty)}$	$\lambda\sqrt{\pi/2}$	$\lambda^2\left(2-\frac{\pi}{2}\right)$		