Statistics for Data Science

MSc Data Science WiSe 2019/20

Prof. Dr. Dirk Ostwald

(4) Random variable transformations

Bibliographic remarks

The majority of the presented material closely follows DeGroot and Schervish (2012, Sections 3.8 - 3.9). The results on the combinations and transformations of Gaussian variables follows Casella and Berger (2002, Section 5.3)

Random variable transformations

- Transformations of discrete random variables
- Direct calculation for continuous random variables
- The probability integral transform
- The univariate probability density function transform
- The multivariate probability density function transform
- Linear combinations of random variables
- Combinations and transformations of Gaussian variables

Random variable transformations

- **Transformations of discrete random variables**
- Direct calculation for continuous random variables
- The probability integral transform
- The univariate probability density function transform
- The multivariate probability density function transform
- Linear combinations of random variables
- Combinations and transformations of Gaussian variables

Theorem (Transformations of discrete random variables)

Let X be a discrete random variable with PMF p_X and let $Y = f(X)$ for some function f on the outcome space of X. Then the PMF of the random variable Y is given by

$$p_Y : \mathcal{Y} \to [0, 1], y \mapsto p_Y(y) := \mathbb{P}(Y = y) = \mathbb{P}(f(X) = y) = \sum_{\{x \mid f(x) = y\}} p_X(x). \tag{1}$$

Example (Discrete uniform distribution)

Let X be a discrete random variable with a finite outcome set $\mathcal{X}$ and probability mass function

$$p : \mathcal{X} \to \mathbb{R}_{\geq 0}, x \mapsto p(x) := \frac{1}{|\mathcal{X}|}. \tag{2}$$

Then X is said to be distributed according to a *discrete uniform distribution*, for which we write $X \sim U(|\mathcal{X}|)$. We abbreviate the PMF of a discrete uniform random variable by

$$U(x; |\mathcal{X}|) := \frac{1}{|\mathcal{X}|}. \tag{3}$$

Example (Transformation of a discrete random variables)

Let X have the discrete uniform distribution on $\mathbb{N}_9$, i.e.,

$$p_X : \mathbb{N}_9 \to [0,1], x \mapsto p_X(x) = \frac{1}{|\mathbb{N}_9|} = \frac{1}{9}. \tag{4}$$

Let $Y = f(X)$ with

$$f : \mathbb{N}_9 \to \mathbb{N}_0^4, x \mapsto f(x) := |x - 5|. \tag{5}$$

We then have

$$p_Y : \mathbb{N}_4^0 \to [0,1],$$

$$\begin{cases} 0 \mapsto p_Y(0) = \mathbb{P}(f(X) = 0) = \sum_{\{x|f(x)=0\}} p_X(x) = p_X(5) & = \frac{1}{9} \\ 1 \mapsto p_Y(1) = \mathbb{P}(f(X) = 1) = \sum_{\{x|f(x)=1\}} p_X(x) = p_X(4) + p_X(6) & = \frac{2}{9} \\ 2 \mapsto p_Y(2) = \mathbb{P}(f(X) = 2) = \sum_{\{x|f(x)=2\}} p_X(x) = p_X(3) + p_X(7) & = \frac{2}{9} \\ 3 \mapsto p_Y(3) = \mathbb{P}(f(X) = 3) = \sum_{\{x|f(x)=3\}} p_X(x) = p_X(2) + p_X(8) & = \frac{2}{9} \\ 4 \mapsto p_Y(4) = \mathbb{P}(f(X) = 4) = \sum_{\{x|f(x)=4\}} p_X(x) = p_X(1) + p_X(9) & = \frac{2}{9} \end{cases} \tag{6}$$

Random variable transformations

- Transformations of discrete random variables
- **Direct calculation for continuous random variables**
- The probability integral transform
- The univariate probability density function transform
- The multivariate probability density function transform
- Linear combinations of random variables
- Combinations and transformations of Gaussian variables

Theorem (Direct calculation for continuous random variables)

Let X be a continuous random variable with PDF p_X and let $Y = f(X)$ for some function f on the outcome space of X. Then the CDF of the random variable Y is given by

$$p_Y : \mathbb{R} \to [0,1], y \mapsto p_Y(y) := \mathbb{P}(Y \leq y) = \mathbb{P}(f(X) \leq y) = \int_{\{x|f(x)\leq y\}} p_X(x)\, dx. \quad (7)$$

If Y is a continuous random variable, then its PDF can be obtained as

$$p_Y : \mathbb{R} \to \mathbb{R}_{\geq 0}, y \mapsto p_Y(y) = \frac{d}{dy} P_Y(y). \quad (8)$$

Example (Continuous uniform distribution)

Let X be a continuous random variable with probability density function

$$p : \mathbb{R} \to \mathbb{R}_{\geq 0}, x \mapsto p(x) := \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & x \notin [a, b]. \end{cases} \tag{9}$$

Then X is said to be distributed according to a *continuous uniform distribution* with parameters a and b, for which we write $X \sim U(a, b)$. We abbreviate the PDF of a continuous uniform random variable by

$$U(x; a, b) := \frac{1}{b-a}. \tag{10}$$

Note that the PDF and CDF of $X \sim U(0, 1)$ are given by

$$p_X : \mathbb{R} \to \mathbb{R}_{\geq 0}, x \mapsto p_X(x) = \begin{cases} 0, & x < 0 \\ 1, & 0 \leq x \leq 1 \\ 0, & x > 1. \end{cases} \quad P_X : \mathbb{R} \to \mathbb{R}_{\geq 0}, x \mapsto P_X(x) := \begin{cases} 0, & x < 0 \\ x, & 0 \leq x \leq 1 \\ 1, & x > 1. \end{cases}, \tag{11}$$

respectively

Example (Direct calculation for a continuous random variable)

Let X have the uniform distribution on the interval $[-1, 1]$, i.e. X has PDF p_X with

$$p_X : [-1, 1] \to \mathbb{R}_{\geq 0}, x \mapsto p_X(x) = \frac{1}{2}. \tag{12}$$

Let $Y = f(X)$ with

$$f : [-1, 1] \to [0, 1], x \mapsto f(x) := x^2. \tag{13}$$

Then the CDF of Y evaluates to

$$\begin{aligned} P_Y : [0,1] \to [0,1], y \mapsto P_X(Y) :=& \\ &= \mathbb{P}(Y \leq y) = \mathbb{P}(X^2 \leq y) = \mathbb{P}\left(-y^{1/2} \leq X \leq y^{1/2}\right) = \mathbb{P}\left(X \in [-y^{1/2}, y^{1/2}]\right) \\ &= \int_{-y^{1/2}}^{y^{1/2}} p_X(x)\, dx = \int_{-y^{1/2}}^{y^{1/2}} \frac{1}{2}\, dx = \frac{1}{2} \int_{-y^{1/2}}^{y^{1/2}} 1\, dx = \frac{1}{2} x|_{y^{-1/2}}^{y^{1/2}} \\ &= \frac{1}{2}\left(y^{1/2} - (-y^{1/2})\right) = \frac{1}{2}\left(2y^{1/2}\right) = y^{1/2}. \end{aligned} \tag{14}$$

Moreover, for $y \in]0, 1[$, the PDF of Y evaluates to

$$p_Y :]0, 1[\to \mathbb{R}_{\geq 0}, y \mapsto p_Y(y) = \frac{d}{dy} P_Y(y) = \frac{1}{2} y^{-1/2}. \tag{15}$$

Random variable transformations

- Transformations of discrete random variables
- Direct calculation for continuous random variables
- **The probability integral transform**
- The univariate probability density function transform
- The multivariate probability density function transform
- Linear combinations of random variables
- Combinations and transformations of Gaussian variables

Theorem (The probability integral transform)

Let X be a continuous random variable with CDF P_X and let

$$Y = P_X(X) \tag{16}$$

be the *probability integral transform*. Then the distribution of Y is the continuous uniform distribution on the interval $[0,1], Y \sim U(0,1)$.

Similarly, let Y have the uniform distribution on the interval $[0,1]$ and let P_X^{-1} be the inverse of a continuous CDF P_X. Then

$$X = P_X^{-1}(Y) \tag{17}$$

has CDF P_X.

Remarks

- Pseudo-random number generators of samples from the uniform distribution can be used to generate samples from arbitrary distributions.
- Let Y have the uniform distribution on $[0,1]$ and let P_X^{-1} denote an inverse CDF. Then $X := P_X^{-1}(Y)$ has CDF P_X.
- Equivalently, let $Y_1, ..., Y_n \sim U(0,1)$. Then $P_X^{-1}(Y_1), ..., P_X^{-1}(Y_n)$ will appear to form an i.i.d. sample from X.

Proof

We first note that for a continuous variable X with P_X, the quantile function of X corresponds to the inverse function of P_X, which we denote by P_X^{-1}. We next note that because P_X is a CDF, we have $0 \le P_X(x) \le 1$ for $x \in \mathbb{R}$. Thus, $\mathbb{P}(Y < 0) = \mathbb{P}(Y > 1) = 0$ and $\mathbb{P}(Y \le 1) = 1 - \mathbb{P}(Y > 1) = 1$. We next consider the CDF P_Y of Y. We have

$$
\begin{aligned}
P_Y(y) &= \mathbb{P}(Y \le y) \\
&= \mathbb{P}\left(P_X(X) \le y\right) \\
&= \mathbb{P}\left(P_X^{-1}\left(P_X(X)\right) \le P_X^{-1}(y)\right) \\
&= \mathbb{P}\left(X \le P_X^{-1}(y)\right) \\
&= P_X\left(P_X^{-1}(y)\right) \\
&= y.
\end{aligned}
\tag{18}
$$

We have thus seen that

$$
P_Y : \mathbb{R} \to \mathbb{R}_{\ge 0} y \mapsto P_Y(y) := \begin{cases} 0, & y < 0 \\ y, & 0 \le y \le 1 \\ 1, & y > 1. \end{cases} \tag{19}
$$

Thus, $Y \sim U(0,1)$. Finally, with

$$
Y = P_X(X) \Leftrightarrow P_X^{-1}(Y) = P_X^{-1}(P_X(X)) \Leftrightarrow X = P_X^{-1}(Y) \tag{20}
$$

the second part of the theorem follows immediately. □

Random variable transformations

- Transformations of discrete random variables
- Direct calculation for continuous random variables
- The probability integral transform
- **The univariate probability density function transform**
- The multivariate probability density function transform
- Linear combinations of random variables
- Combinations and transformations of Gaussian variables

Theorem (The univariate probability density function transform)

Let X be a random variable with PDF p_X and for which $\mathbb{P}(]a,b[) = 1$, where a and/or b are either finite or infinite. Let $Y = f(X)$, where f is differentiable and bijective for $]a,b[$. Let $f(]a,b[$ be the image of $]a,b[$ under f. Finally, let $f^{-1}(y)$ denote the inverse of $f(x)$ for $y \in f(]a,b[)$ and let $f'(x)$ denote the first derivative of f at x.

Then the PDF of Y is given by

$$p_Y : \mathbb{R} \to \mathbb{R}_{\geq 0}, y \mapsto p_Y(y) := \begin{cases} \frac{1}{|f'(f^{-1}(y))|} p_X\left(f^{-1}(y)\right) & \text{for } y \in f(]a,b[) \\ 0 & \text{for } y \in \mathbb{R} \setminus f(]a,b[). \end{cases} \tag{21}$$

Proof

We first note that because f is a differentiable bijective function on the open interval $]a,b[$ it is either strictly increasing or strictly decreasing. Assume first that f is increasing on $]a,b[$. Then f^{-1} is also increasing for all $y \in f(]a,b[)$ and

$$P_Y(y) = \mathbb{P}(Y \le y) = \mathbb{P}(f(X) \le y) = \mathbb{P}\left(f^{-1}(f(X)) \le f^{-1}(y)\right) = \mathbb{P}\left(X \le f^{-1}(y)\right) = P_X\left(f^{-1}(y)\right). \tag{22}$$

P_Y is thus differentiable at all y where both f^{-1} and P_X is differentiable at $f^{-1}(y)$. With the chain rule of differentiation and the inverse function theorem $(f^{-1}(x))' = 1/f'(f^{-1}(x))$, it thus follows that the PDF p_Y evaluates to

$$p_Y(y) = \frac{d}{dy}P_Y(y) = \frac{d}{dy}P_X\left(f^{-1}(y)\right) = p_X\left(f^{-1}(y)\right)\frac{d}{dy}f^{-1}(y) = \frac{1}{f'\left(f^{-1}(y)\right)}p_X\left(f^{-1}(y)\right), \tag{23}$$

Because f^{-1} is increasing, $d/dy(f^{-1}(y))$ is positive, and the theorem holds. Similarly, if f is decreasing on $]a,b[$, then f^{-1} is also decreasing for all $y \in f(]a,b[)$ and for each $y \in f(]a,b[)$ and

$$P_Y(y) = \mathbb{P}(f(X) \le y) = \mathbb{P}\left(f^{-1}(f(X)) \ge f^{-1}(y)\right) = \mathbb{P}\left(X \ge f^{-1}(y)\right) = 1 - P_X\left(f^{-1}(y)\right), \tag{24}$$

With the chain rule of differentiation and the inverse function theorem, it thus follows that the PDF p_Y evaluates to

$$p_Y(y) = \frac{d}{dy}(1 - P_Y(y)) = -\frac{d}{dy}P_X\left(f^{-1}(y)\right) = -p_X\left(f^{-1}(y)\right)\frac{d}{dy}f^{-1}(y) = -\frac{1}{f'\left(f^{-1}(y)\right)}p_X\left(f^{-1}(y)\right). \tag{25}$$

Since f^{-1} is strictly decreasing, $d/dy(f^{-1}(y))$ is negative, such that $-d/dy(f^{-1}(y))$ equals $|d/dy(f^{-1}(y))|$ and the theorem holds.

□

Theorem (The univariate PDF transform for linear functions)

Let X be a random variable with PDF p_X and let $Y = f(X)$ with

$$f(x) := ax + b \text{ for } a \neq 0. \tag{26}$$

Then the PDF of Y is given by

$$p_Y : \mathbb{R} \to \mathbb{R}_{\geq 0}, y \mapsto p_Y(y) := \frac{1}{|a|} p_X\left(\frac{y-b}{a}\right). \tag{27}$$

<u>Proof</u>

We first note that

$$f^{-1} : \mathbb{R} \to \mathbb{R}, y \mapsto f^{-1}(y) = \frac{y-b}{a} \tag{28}$$

because then $f \circ f^{-1} = \mathrm{id}_{\mathbb{R}}$ as

$$f(f^{-1}(x)) = a\left(\frac{x-b}{a}\right) + b = x - b + b = x \text{ for all } x \in \mathbb{R} \tag{29}$$

We next note that

$$f' : \mathbb{R} \to \mathbb{R}, x \mapsto f'(x) = \frac{d}{dx}(ax+b) = a. \tag{30}$$

We thus have

$$\begin{aligned} p_Y : \mathbb{R} \to \mathbb{R}_{\geq 0}, y \mapsto p_Y(y) &= \frac{1}{|f'\left(f^{-1}(y)|\right)} p_X\left(f^{-1}(y)\right) \\ &= \frac{1}{|a|} p_X\left(\frac{y-b}{a}\right). \end{aligned} \tag{31}$$

□

Example (Linear transformation of a standard normal variable)

Let $X \sim N(0,1)$ and $Y = f(X)$ with $f(x) := ax + b$. Then $Y \sim N(b, a^2)$.

Proof

We first note that $f^{-1}(y) = \frac{y-b}{a}$ and $f'(x) = a$. With the univariate PDF transform for linear functions, we then have for the PDF of Y

$$
\begin{aligned}
p_Y(y) &= \frac{1}{|a|} N\left(\frac{y-b}{a}; 0, 1\right) \\
&= \frac{1}{|a|}\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y-b}{a}\right)^2\right) \\
&= \frac{1}{\sqrt{a^2}}\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2a^2}(y-b)^2\right) \\
&= \frac{1}{\sqrt{2\pi a^2}} \exp\left(-\frac{1}{2a^2}(y-b)^2\right) \\
&= N\left(y; b, a^2\right)
\end{aligned}
\tag{32}
$$

and thus $Y \sim N\left(b, a^2\right)$.

□

Example (Z-transformation)

Let $X \sim N(\mu, \sigma^2)$ and $Y = f(X)$ with $f(x) := \frac{x-\mu}{\sigma}$. Then $Y \sim N(0,1)$.

Proof

We first note that $f^{-1}(y) = \sigma y + \mu$ and $f'(x) = \frac{1}{\sigma}$. With the univariate PDF transform for linear functions, we then have for the PDF of Y

$$
\begin{aligned}
p_Y(y) &= \frac{1}{|1/\sigma|} N\left(\sigma y + \mu; \mu, \sigma^2\right) \\
&= \frac{1}{1/\sqrt{\sigma^2}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\left(\sigma y + \mu - \mu\right)^2\right) \\
&= \frac{\sqrt{\sigma^2}}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\sigma^2 y^2\right) \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right) \\
&= N(y; 0, 1)
\end{aligned}
\tag{33}
$$

and thus $Y \sim N(0,1)$.

□

Random variable transformations

- Transformations of discrete random variables
- Direct calculation for continuous random variables
- The probability integral transform
- The univariate probability density function transform
- **The multivariate probability density function transform**
- Linear combinations of random variables
- Combinations and transformations of Gaussian variables

Theorem (The multivariate probability density function transform)

Let X be an n-dimensional random vector with PDF p_X and let $Y = f(X)$ be an n-dimensional random vector, where $f : \mathbb{R}^n \to \mathbb{R}^n$ is differentiable and bijective. Let $f^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ denote the inverse of f. Let further

$$J^f(x) = \left(\frac{\partial}{\partial x_j} f_i(x)\right)_{i,j=1,\dots,n} \in \mathbb{R}^{n\times n} \tag{34}$$

denote the Jacobian of f at $x \in \mathbb{R}$, let $|J^f(x)|$ denote its determinant, and assume that $|J^f(x)| \neq 0$ for all $x \in \mathbb{R}^n$. Then the PDF of Y is given by

$$p_Y : \mathbb{R}^n \to \mathbb{R}_{\geq 0}, y \mapsto p_Y(y) := \begin{cases} \frac{1}{|J^f(f^{-1}(y))|} p_X\left(f^{-1}(y)\right) & \text{for } y \in f(\mathbb{R}^n) \\ 0 & \text{for } y \in \mathbb{R}^n \setminus f(\mathbb{R}^n). \end{cases} \tag{35}$$

Remarks

- A straight-forward generalization of the univariate case.
- Proof omitted.

Theorem (The multivariate PDF transform for linear functions)

Let X be a random vector with PDF p_X. Let $Y = f(X)$ with

$$f(x) = Ax, A \in \mathbb{R}^{n \times n} \text{ and nonsingular.} \tag{36}$$

Then the PDF of Y is given by

$$p_Y : \mathbb{R}^n \to \mathbb{R}_{\geq 0} y \mapsto p_Y(y) = \frac{1}{|A|} p_X(A^{-1}y) \tag{37}$$

where $|A|$ and A^{-1} denote the determinant and the inverse of A, respectively.

Remarks

- A straight-forward application the multivariate PDF transform theorem.
- Central for classical and Bayesian linear Gaussian models.

Proof

We first show that

$$f^{-1} : \mathbb{R}^n \to \mathbb{R}^n, y \mapsto f^{-1}(y) := A^{-1}y. \tag{38}$$

To this end, we note that

$$f^{-1}(f(x)) = A^{-1}Ax = x = \mathrm{id}_{\mathbb{R}^n}(x). \tag{39}$$

We next show that

$$J^f\left(f^{-1}(y)\right) = A. \tag{40}$$

To this end, we first note that

$$f_i(x) = \sum_{j=1}^{n} a_{ij}x_j. \tag{41}$$

Thus

$$\begin{aligned}
J^f(x) &= \left(\frac{\partial}{\partial x_j} f_i(x)\right)_{i,j=1,\ldots,n} &(42)\\
&= \left(\sum_{j=1}^{n} \frac{\partial}{\partial x_j} a_{ij}x_j\right)_{i,j=1,\ldots,n} &(43)\\
&= (a_{ij})_{i,j=1,\ldots,n} &(44)\\
&= A \in \mathbb{R}^{n\times n}, &(45)
\end{aligned}$$

which is constant.

□

Example (Linear transformations of multivariate Gaussians)

If

- X is a random vector distributed according to an n-variate Gaussian distribution with probability density function $N(x; \mu_x, \Sigma_x)$ for $x, \mu_x \in \mathbb{R}^n, \Sigma_x \in \mathbb{R}^{n \times n}$ p.d., and
- $A \in \mathbb{R}^{n \times n}$ is a matrix of full column-rank

then the random vector

$$Y := AX \tag{46}$$

is distributed according to an n-variate Gaussian distribution with probability density function $N(y; \mu_y, \Sigma_y)$ with $y, \mu_y \in \mathbb{R}^m$ and $\Sigma_y \in \mathbb{R}^{n \times n}$, and in particular

$$\mu_y := A\mu_x \text{ and } \Sigma_y := A\Sigma_x A^T. \tag{47}$$

Example (Linear transformations of multivariate Gaussians cont.)

Proof

With the multivariate PDF transform theorem for linear functions, we first note that we have for the normalization term

$$|A|^{-1}(2\pi)^{-\frac{n}{2}}|\Sigma|^{-\frac{1}{2}} = |A|^{-\frac{1}{2}}|\Sigma|^{-\frac{1}{2}}|A^T|^{-\frac{1}{2}} = |A\Sigma A^T|^{-\frac{1}{2}} \tag{48}$$

We next note that for the the argument of the exponential term, we have

$$\begin{aligned}
&-\frac{1}{2}(A^{-1}y-\mu)^T\Sigma^{-1}(A^{-1}y-\mu) \\
&\quad= -\frac{1}{2}(yA^{-1}\Sigma^{-1}A^{-1}y - 2yA^{-1}\Sigma^{-1}\mu + \mu^T\Sigma^{-1}\mu) \\
&\quad= -\frac{1}{2}(yA^{-1}\Sigma^{-1}A^{-1}y - 2yA^{-1}\Sigma^{-1}A^{-1}A\mu + \mu^T\Sigma^{-1}\mu) \\
&\quad= -\frac{1}{2}(yA^{-1}\Sigma^{-1}A^{-1}y - 2yA^{-1}\Sigma^{-1}A^{-1}A\mu + (A^{-1}A)^T\mu^T\Sigma^{-1}A^{-1}A\mu) \\
&\quad= -\frac{1}{2}(y-A\mu)^T(A\Sigma A^T)^{-1}(y-A\mu),
\end{aligned} \tag{49}$$

which completes the proof. □

Random variable transformations

- Transformations of discrete random variables
- Direct calculation for continuous random variables
- The probability integral transform
- The univariate probability density function transform
- The multivariate probability density function transform
- **Linear combinations of random variables**
- Combinations and transformations of Gaussian variables

Theorem (Convolution of continuous random variables)

Let X and Y be two independent random variable with PDFs p_X and p_Y, respectively. Let

$$Z := X + Y \tag{50}$$

be the sum of X and Y. Then a PDF of Z is given by the *convolution* of the PDFs of X and Y,

$$p_Z : \mathbb{R} \to \mathbb{R}_{\geq 0}, z \mapsto p_Z(z) = (p_X * p_Y)(z), \tag{51}$$

defined as

$$(p_X * p_Y)(z) := \int_{-\infty}^{\infty} p_X(z - \xi) p_Y(\xi)\, d\xi = \int_{-\infty}^{\infty} p_X(\xi) p_Y(z - \xi), d\xi. \tag{52}$$

Remark

- For a proof, see DeGroot and Schervish (2012, p. 178 - 179).
- Fundamental for classical parametric confidence intervals and statistical testing.
- Fundamental for nonparametric kernel density estimation.

Random variable transformations

- Transformations of discrete random variables
- Direct calculation for continuous random variables
- The probability integral transform
- The univariate probability density function transform
- The multivariate probability density function transform
- Linear combinations of random variables
- **Combinations and transformations of Gaussian variables**

Theorem (Distribution of Gaussian sample means and variances)

Let $X_1, ..., X_n \sim N(\mu, \sigma^2)$ and let

$$\bar{X} := \frac{1}{n}\sum_{i=1}^{n} X_i \text{ and } S^2 := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \tag{53}$$

denote the *sample mean* and *sample variance* of $X_1, ..., X_n$. Then

1. $\bar{X}$ and S^2 are independent,
2. $\bar{X} \sim N(\mu, \sigma^2/n)$, and
3. $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$.

Remark

- For a proof, see Casella and Berger (2002, pp. 218 - 219)

Example (Chi-squared distribution)

Let X be a continuous random variable with probability density function

$$p : \mathbb{R}_{>0} \to \mathbb{R}_{>0}, x \mapsto p(x) := \frac{1}{\Gamma(d/2)2^{d/2}} x^{(d/2)-1} \exp\left(-\frac{1}{2}x\right), \tag{54}$$

where Γ denotes the Gamma function. Then X is said to be distributed according to a *chi-squared distribution* with d degrees of freedom, for which we write $X \sim \chi^2(d)$. We abbreviate the PDF of a chi-squared random variable by

$$\chi^2(x; d) := \frac{1}{\Gamma(d/2)2^{d/2}} x^{(d/2)-1} \exp\left(-\frac{1}{2}x\right). \tag{55}$$

Theorem (Properties of chi-squared random variables)

1. If $Z \sim N(0,1)$, then $Z^2 \sim \chi^2(1)$.
2. If $X_1, ..., X_n$ are independent chi-squared random variables with degrees of freedom $d_1, ..., d_n$, then $\sum_{i=1}^n X_i^2 \sim \chi^2\left(\sum_{i=1}^n d_i\right)$

Remarks

- A chi-squared variable is a squared standard normal variable.
- The sum of independent chi-squared variables has a chi-squared distribution with degrees of freedom corresponding to the sum of the individual degrees of freedom.
- For a proof, see Casella and Berger (2002, p. 219)

Definition (t-distribution)

Let $X_1, ..., X_n \sim N(\mu, \sigma^2)$ and let

$$T := \sqrt{n}\frac{\bar{X} - \mu}{S}, \text{ where } S := \sqrt{S^2}, \tag{56}$$

denote the *T-statistic*. Then $T \sim \mathrm{T}(n-1)$, i.e. T has a *t distribution with $n-1$ degrees of freedom*. Equivalently, let X be a continuous random variable with probability density function

$$p : \mathbb{R} \to \mathbb{R}_{\geq 0}, x \mapsto p(x) := \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} \frac{1}{(d\pi)^{1/2}} \left(\left(1 + \frac{x^2}{d}\right)^{(d+1)/2}\right)^{-1}. \tag{57}$$

Then X is said to be distributed according to a *t distribution with d degrees of freedom*, for which we write $X\mathrm{T}(d)$. We abbreviate the PDF of a t distribution with d degrees of freedom by

$$\mathrm{t}(t; d) := \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} \frac{1}{(d\pi)^{1/2}} \left(\left(1 + \frac{x^2}{d}\right)^{(d+1)/2}\right)^{-1}. \tag{58}$$

Remark

- For a proof, see Casella and Berger (2002, p. 223 - 224)