Computation of Statistical Distributions

Documentation of the Program ELV

Second Edition

****

by Leo Knüsel, University of Munich

This documentation is the English translation (with some minor improvements) of

Computergestützte Berechnung statistischer Verteilungen
von Leo Knüsel, Universität München
Verlag: R. Oldenbourg, München, Wien, 1989

(translated by Leo Knüsel, January 1998).

Second Edition: 2003
This version can be accessed under

`http://www.stat.uni-muenchen.de/~knuesel`

with kindest permission of Oldenbourg, Munich.

Foreword

This brochure is the documentation of the program ELV (for elementary distributions, in German: **EL**ementare **V**erteilungen). The program ELV computes upper and lower tail probabilities as well as upper and lower quantiles of some important statistical distributions, and so one can do without extensive statistical tables for the distributions considered. The author claims that all results given by the program are correct; for all but the noncentral distributions this is true for probabilities as small as 10^{-100}. Of course I am very indebted to any reader pointing out some possible errors, as claim and reality may disagree also with ELV.

Bernhard Bablok (1988) has made some important contributions to the computation of the noncentral distributions in his diploma thesis at the University of Munich; he has discovered several incorrect formulas in the statistical literature, and he has developed the programs for computing the noncentral distributions in ELV. Furthermore I wish to express my thanks to Bernhard Bablok for typing this brochure in PC-TEX and for drawing the figures. I also thank to Rainer Würländer who completed the final version of this documentation.

Munich, April 1989 Leo Knüsel

Address: Leo Knüsel
Department of Statistics
University of Munich
Geschwister-Scholl-Platz 1
80539 Munich, Germany

e-mail: knuesel@stat.uni-muenchen.de

Foreword to Second Edition

The program ELV now computes the noncentral distributions with the same accuracy standard as with the other distributions (upper and lower tail probabilities with six significant digits for probabilities as small as 10^{-100}). The noncentrality parameter with the Chi-Square and F distribution has been redefined in order to comply with most statistical packages. I thank to Jerry W. Lewis for pointing out some inconsistencies in the documentation.

Munich, October 2003 Leo Knüsel

Contents

1 Purpose of the Program

The program ELV computes probabilities and quantiles of some elementary statistical distributions. The upper and lower tail probabilities of all distributions are computed with six significant digits for probabilities as small as 10^{-100}; probabilities smaller than 10^{-100} are rounded to zero. Upper and lower quantiles are computed for all distributions 1–9 (see below) for tail probabilities P with $10^{-12} \leq P \leq \frac{1}{2}$.

2 Using the Program

When starting the program the following main menu appears:

```
Main Menu
  1    Standard normal distribution
  2    Gamma distribution
  3    Chi-square distribution
  4    Beta distribution
  5    F distribution
  6    t distribution (Student's distribution)
  7    Poisson distribution
  8    Binomial distribution
  9    Hypergeometric distribution
-1,...,-9:  Quantiles of the distributions 1,...,9
12,...,16:  Noncentral version of the distributions 2,...,6
Choose a distribution by entering the corresponding number:
```

If I want to compute e.g. quantiles of the standard normal distribution I enter −1 and press the ENTER key. The program answers:

```
Quantiles of the standard normal distribution
Enter P:
```

I enter the value 0.05, and after pressing the ENTER key I obtain the answer:

```
Input:  P = .05                 Output:  zr = 1.64485
Quantiles of the standard normal distribution
Enter P:
```

Now I can compute some other quantiles of the standard normal distribution by entering the corresponding P-values, or I can switch to some other distribution by pressing the Esc key (Escape key) or the F6 key (F6 = CTRL+Z = EOF = end of file) which brings me back to the main menu. If I want to leave the program I press again the Esc key (or the F6 key).

Notation:

Instead of	z_ℓ	z_r	x_ℓ	x_r	k_ℓ	k_r	n_1	n_2	ϑ	λ	δ
ELV writes	`zl`	`zr`	`xl`	`xr`	`kl`	`kr`	`n1`	`n2`	`theta`	`lambda`	`delta`

3 Entering Data

The input prompt of the program informs about the number and type (integer or real) of the input data. If several input values are required these values can be entered on the same line separated by commas or blanks; if the ENTER key is pressed before entering all necessary parameters the missing data can be entered on a new line (list directed read in FORTRAN).

Examples:

With the binomial distribution (number 8) the input prompt is:

```
Enter k, n, theta:
```

k and n are integer parameters but ϑ is real. The following table gives for some entered data of the user the corresponding values that the program processes. With $\hookleftarrow$ we denote the ENTER key.

	Entered data	Data processed by ELV
1)	2 60 0.25 $\hookleftarrow$	$k = 2$ $n = 60$ $\vartheta = 0.25$
2)	$2, 60, 0.25$ $\hookleftarrow$	$k = 2$ $n = 60$ $\vartheta = 0.25$ Blanks and commas are separators.
3)	2 $\hookleftarrow$ 60$\hookleftarrow$ .25$\hookleftarrow$	The program waits for further data. The program waits for further data. $k = 2$ $n = 60$ $\vartheta = 0.25$
4)	5, , , $\hookleftarrow$	$k = 5$ $n = 60$ $\vartheta = 0.25$ If no new data are entered the corresponding old values are used.
5)	, 70, , $\hookleftarrow$	$k = 5$ $n = 70$ $\vartheta = 0.25$
6)	, , 0.3 $\hookleftarrow$	$k = 5$ $n = 70$ $\vartheta = 0.3$
7)	2∗100, , $\hookleftarrow$	$k = 100$ $n = 100$ $\vartheta = 0.3$ Repeating factors enable multiple entry of the same value.
8)	1e2 2.5e3 1e–2 $\hookleftarrow$	$k = 100$ $n = 2500$ $\vartheta = 0.01$ Scientific notation is possible.
9)	$100, 1\ 000, 0.4$ $\hookleftarrow$	$k = 100$ $n = 1$ $\vartheta = 0.0$ Blank acts as a separator!
10)	100 1000 $0, 4$ $\hookleftarrow$	$k = 100$ $n = 1000$ $\vartheta = 0.0$ Comma acts as a separator!
11)	5.9 10.7 1 $\hookleftarrow$	$k = 5$ $n = 10$ $\vartheta = 1.0$ Fractional part is truncated with an integer argument.)
12)	k=2 $\hookleftarrow$	Illegal entry.

4 Description of the Parameters

4.1 Standard Normal Distribution

4.1.1 Tail probabilities of the standard normal distribution:

Let $Z \sim N(0,1)$.

Input:	z

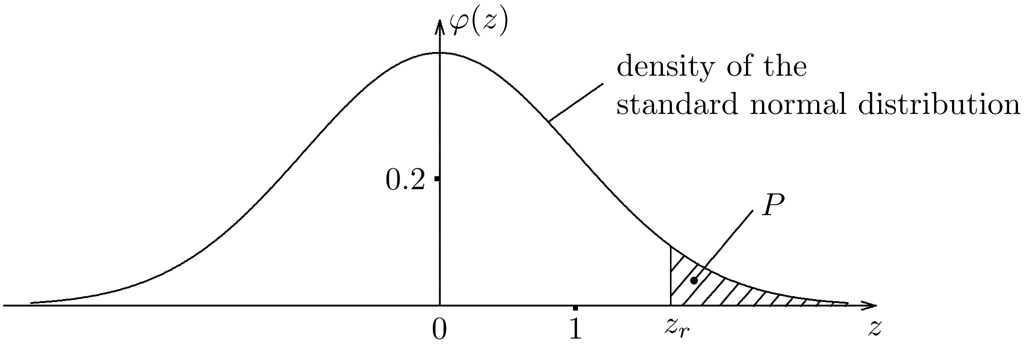
Output:	$I\!P\{Z < z\}$
	$I\!P\{Z > z\}$

4.1.2 Quantiles of the standard normal distribution:

Let $Z \sim N(0,1)$. For given $P \in (0,1)$ we define the right (upper) P-quantile $z_r = z_r(P)$ as:

$$z_r = z_r(P): \quad I\!P\{Z > z_r\} = P$$

Fig. 1: Quantiles of the standard normal distribution



Input:	P

Output:	z_r

Admissible values of P:

$$10^{-100} \le P \le \tfrac{1}{2}.$$

4.2 Gamma Distribution

4.2.1 Tail probabilities of the Gamma distribution:

Let $X \sim Ga(a)$ with $a > 0$.

Input:	x
	a

Output:	$I\!P\{X < x\}$
	$I\!P\{X > x\}$

Admissible values of a and x:

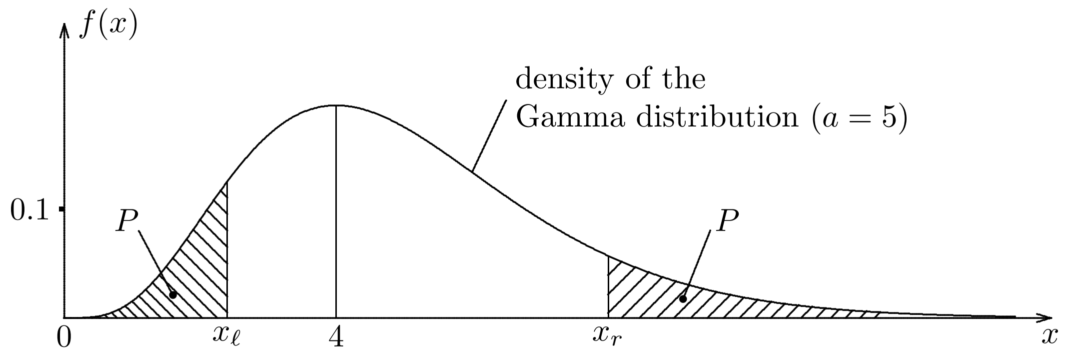
$$0 < a \le 2^{26} \ (\approx 67 \cdot 10^6); \quad x \ge 0.$$

4.2.2 Quantiles of the Gamma distribution:

Let $X \sim Ga(a)$ with $a > 0$. For given $P \in (0,1)$ left and right P-quantiles are defined as:

$$x_\ell = x_\ell(P,a): \ \mathbb{P}\{X < x_\ell\} = P$$
$$x_r = x_r(P,a): \ \mathbb{P}\{X > x_r\} = P$$

Fig. 2: Quantiles of the Gamma distribution



Input:	P
	a

Output:	x_ℓ
	x_r

Admissible values of a and P:

$0 < a \leq 2^{26}; \quad 10^{-12} \leq P \leq \frac{1}{2}.$

4.2.3 Noncentral Gamma distribution:

Let $X \sim Ga(a,\lambda)^{(*)}$ with $a > 0,\ \lambda \geq 0.$

Input:	x
	a
	λ

Output:	$\mathbb{P}\{X < x\}$
	$\mathbb{P}\{X > x\}$

Admissible values of a, λ and x:

$a > 0; \quad \lambda \geq 0; \quad a + \lambda \leq 2^{26}; \quad x \geq 0.$

(*) To keep notation simple we denote the central and noncentral Gamma distribution by $Ga(a)$ and $Ga(a,\lambda)$, respectively. The analogous notation will be used for the other noncentral distributions.

4.3 Chi-square Distribution

4.3.1 Tail probabilities of the chi-square distribution:

Let $X \sim \chi^2(n)$ with $n = 1, 2, \ldots$

Input:	
	x
	n

Output:	
	$\mathbb{P}\{X < x\}$
	$\mathbb{P}\{X > x\}$

Admissible values of n and x:

$1 \leq n \leq 2^{27}$ $(\approx 134 \cdot 10^6)$; $\quad x \geq 0$.

4.3.2 Quantiles of the chi-square distribution:

Let $X \sim \chi^2(n)$ with $n = 1, 2, \ldots$. The quantiles $x_\ell = x_\ell(P, n)$ and $x_r = x_r(P, n)$ are defined as with the Gamma distribution.

Input:	
	P
	n

Output:	
	x_ℓ
	x_r

Admissible values of n and P:

$1 \leq n \leq 2^{27}$; $\quad 10^{-12} \leq P \leq \frac{1}{2}$.

4.3.3 Noncentral chi-square distribution:

Let $X \sim \chi^2(n, \lambda)$ with $n = 1, 2, \ldots,$ $\lambda \geq 0$.

Input:	
	x
	n
	λ

Output:	
	$\mathbb{P}\{X < x\}$
	$\mathbb{P}\{X > x\}$

Admissible values of n, λ and x:

$n \geq 1$; $\quad \lambda \geq 0$; $\quad n + \lambda \leq 2^{27}$; $\quad x \geq 0$.

4.4 Beta Distribution

4.4.1 Tail probabilities of the Beta distribution:

Let $X \sim Be(a,b)$ with $a,b>0$.

Input:	
	x
	a
	b

Output:	
	$\mathbb{P}\{X<x\}$
	$\mathbb{P}\{X>x\}$

Admissible values of a, b and x:

$0 < a, b \le 2^{26}; \quad 0 \le x \le 1.$

4.4.2 Quantiles of the Beta distribution:

Let $X \sim Be(a,b)$ with $a,b>0$. The quantiles $x_\ell = x_\ell(P,a,b)$ and $x_r = x_r(P,a,b)$ are defined as with the Gamma distribution.

Input:	
	P
	a
	b

Output:	
	x_ℓ
	x_r

Admissible values of a, b and P:

$0 < a, b \le 2^{26}; \quad 10^{-12} \le P \le \frac{1}{2}.$

4.4.3 Noncentral Beta distribution:

Let $X \sim Be(a,b,\lambda)$ with $a,b>0$, $\lambda \ge 0$.

Input:	
	x
	a
	b
	λ

Output:	
	$\mathbb{P}\{X<x\}$
	$\mathbb{P}\{X>x\}$

Admissible values of a, b, λ and x:

$a, b > 0; \quad \lambda \ge 0; \quad a+\lambda \le 2^{26}; \quad b \le 2^{26}; \quad 0 \le x \le 1.$

4.5 F Distribution

4.5.1 Tail probabilities of the F distribution:

Let $X \sim F(n_1, n_2)$ with $n_1, n_2 = 1, 2, \ldots$

Input:	
	x
	n_1
	n_2

Output:	
	$\mathbb{P}\{X < x\}$
	$\mathbb{P}\{X > x\}$

Admissible values of n_1, n_2 and x:

$$1 \leq n_1, n_2 \leq 2^{27}; \quad x \geq 0.$$

4.5.2 Quantiles of the F distribution:

Let $X \sim F(n_1, n_2)$ with $n_1, n_2 = 1, 2, \ldots$.
The quantiles $x_\ell = x_\ell(P, n_1, n_2)$ and $x_r = x_r(P, n_1, n_2)$ are defined as with the Gamma distribution.

Input:	
	P
	n_1
	n_2

Output:	
	x_ℓ
	x_r

Admissible values of n_1, n_2 and P:

$$1 \leq n_1, n_2 \leq 2^{27}; \quad 10^{-12} \leq P \leq \tfrac{1}{2}.$$

4.5.3 Noncentral F distribution:

Let $X \sim F(n_1, n_2, \lambda)$ with $n_1, n_2 = 1, 2, \ldots,\ \lambda \geq 0$.

Input:	
	x
	n_1
	n_2
	λ

Output:	
	$\mathbb{P}\{X < x\}$
	$\mathbb{P}\{X > x\}$

Admissible values of n_1, n_2, λ and x:

$$n_1, n_2 \geq 1; \quad \lambda \geq 0; \quad n_1 + \lambda \leq 2^{27}; \quad n_2 \leq 2^{27}; \quad x \geq 0.$$

4.6 t Distribution

4.6.1 Tail probabilities of the t distribution:

Let $X \sim t(n)$ with $n = 1, 2, \ldots$

Input:	
	x
	n

Output:	
	$\mathbb{P}\{X < x\}$
	$\mathbb{P}\{X > x\}$

Admissible values of n:

$1 \leq n \leq 2^{27}$.

4.6.2 Quantiles of the t distribution:

Let $X \sim t(n)$ with $n = 1, 2, \ldots$. The (central) t distribution is symmetric and the quantile $x_r = x_r(P, n)$ is defined as with the standard normal distribution.

Input:	
	P
	n

Output:	
	x_r

Admissible values of n and P:

$1 \leq n \leq 2^{27}; \quad 10^{-12} \leq P \leq \frac{1}{2}$.

4.6.3 Noncentral t distribution:

Let $X \sim t(n, \delta)$ with $n = 1, 2, \ldots,\ \delta \in \mathbb{R}$.

Input:	
	x
	n
	δ

Output:	
	$\mathbb{P}\{X < x\}$
	$\mathbb{P}\{X > x\}$
	$\mathbb{P}\{\lvert X\rvert < x\}$
	$\mathbb{P}\{\lvert X\rvert > x\}$

Admissible values of n and δ:

$1 \leq n \leq 2^{27}; \quad 0 \leq \delta^2 \leq 2^{27}$ (i.e. $|\delta| \leq 11\,585$).

4.7 Poisson Distribution

4.7.1 Tail and point probabilities of the Poisson distribution:

Let $X \sim Po(\lambda)$ with $\lambda \geq 0$.

Input:	
	k
	λ

Output:	
	$\mathbb{P}\{X \leq k\}$
	$\mathbb{P}\{X > k\}$
	$\mathbb{P}\{X = k\}$

Admissible values of λ and k:

$$0 \leq \lambda \leq 2^{26}; \quad k \geq 0$$

4.7.2 Quantiles of the Poisson distributon:

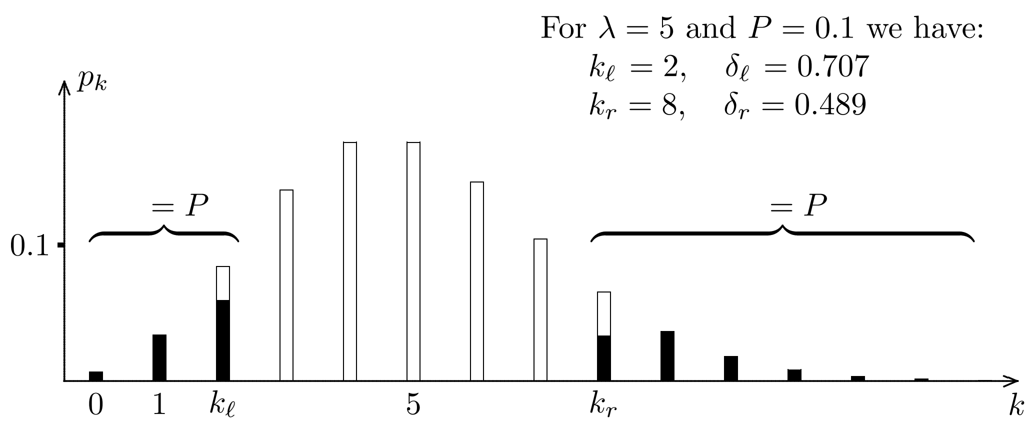
Let $X \sim Po(\lambda)$ with $\lambda \geq 0$. For given $P \in (0,1)$ we define the left P-quantile by the two quantities k_ℓ and δ_ℓ as follows:

$k_\ell = k_\ell(P,\lambda)$ is the largest integer with $\mathbb{P}\{X < k_\ell\} \leq P$.
$\delta_\ell = \delta_\ell(P,\lambda)$: $\mathbb{P}\{X < k_\ell\} + \delta_\ell \cdot \mathbb{P}\{X = k_\ell\} = P$ with $0 \leq \delta_\ell < 1$.

The right P-quantile is defined by the two quantities k_r and δ_r as follows:

$k_r = k_r(P,\lambda)$ is the smallest integer with $\mathbb{P}\{X > k_r\} \leq P$.
$\delta_r = \delta_r(P,\lambda)$: $\mathbb{P}\{X > k_r\} + \delta_r \cdot \mathbb{P}\{X = k_r\} = P$ with $0 \leq \delta_r < 1$.

Fig. 3: Quantiles of the Poisson distribution



Input:	
	P
	λ

Output:	
	$k_\ell \ (\delta_\ell)$
	$k_r \ (\delta_r)$

Admissible values of λ und P:

$$0 \leq \lambda \leq 2^{26}; \quad 10^{-12} \leq P \leq \tfrac{1}{2}.$$

4.8 Binomial Distribution

4.8.1 Tail and point probabilities of the binomial distribution:

Let $X \sim Bi(n, \vartheta)$ with $0 \leq \vartheta \leq 1$.

Input:	
	k
	n
	ϑ

Output:	
	$\mathbb{P}\{X \leq k\}$
	$\mathbb{P}\{X > k\}$
	$\mathbb{P}\{X = k\}$

Admissible values of n, ϑ and k:

$1 \leq n \leq 2^{26}$;

$0 \leq \vartheta \leq 1$;

$0 \leq k \leq n$.

4.8.2 Quantiles of the binomial distribution:

Let $X \sim Bi(n, \vartheta)$ with $0 \leq \vartheta \leq 1$. The quantiles (k_ℓ, δ_ℓ) and (k_r, δ_r) are defined as with the Poisson distribution.

Input:	
	P
	n
	ϑ

Output:	
	$k_\ell \ (\delta_\ell)$
	$k_r \ (\delta_r)$

Admisssible values of n, ϑ and P:

$1 \leq n \leq 2^{26}$;

$0 \leq \vartheta \leq 1$;

$10^{-12} \leq P \leq \frac{1}{2}$.

4.9 Hypergeometric Distribution

4.9.1 Tail and point probabilities of the hypergeometric distribution:

Let $X \sim H(N, M, n)$.

Input:	
	k
	N
	M
	n

Output:	
	$\mathbb{P}\{X \leq k\}$
	$\mathbb{P}\{X > k\}$
	$\mathbb{P}\{X = k\}$

Admissible values of N, M, n and k:

$2 \leq N \leq 2^{26}$;

$0 < n, M < N$;

$0 \leq k \leq M$;

$0 \leq n - k \leq N - M$.

4.9.2 Quantiles of the hypergeometric distribution:

Let $X \sim H(N, M, n)$. The quantiles (k_ℓ, δ_ℓ) and (k_r, δ_r) are defined as with the Poisson distribution.

Input:	
	P
	N
	M
	n

Output:	
	k_ℓ (δ_ℓ)
	k_r (δ_r)

Admissible values of N, M, n und P:

$2 \leq N \leq 2^{26}$;

$0 < n, M < N$;

$10^{-12} \leq P \leq \frac{1}{2}$.

5 Definition of the Distributions

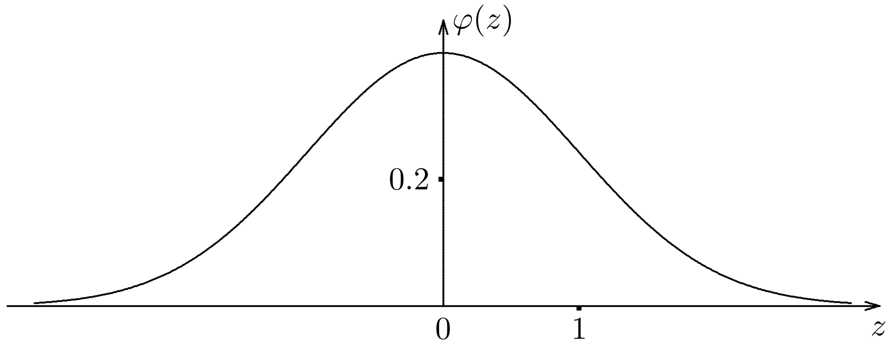
5.1 Standard Normal Distribution

Let $Z \sim N(0,1)$, i.e. let Z be a random variable with a standard normal distribution.

- Density function of Z:

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty.$$

Fig. 4: Density of the standard normal distribution



- Expectation, variance and mode:

$$\mathbb{E}(X) = 0, \quad var(X) = 1, \quad mode(X) = 0.$$

5.2 Gamma Distribution

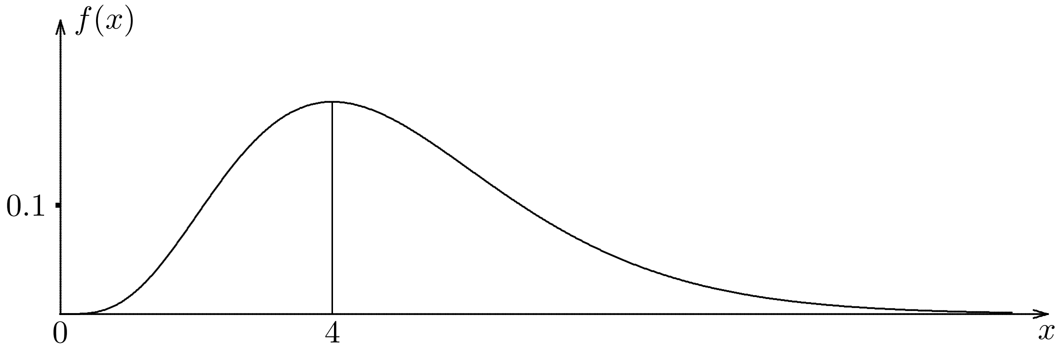
5.2.1 (Central) Gamma distribution

Let $X \sim Ga(a)$, $a > 0$, i.e. let X be a random variable with a Gamma distribution with parameter a.

- Density function of X for $x > 0$:

$$f(x) = \frac{1}{\Gamma(a)} x^{a-1} e^{-x} \quad \text{with} \quad \Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt \quad \text{(Gamma function).}$$

Fig. 5: Density of the Gamma distribution ($a = 5$)



- Expectation, variance and mode:

$$\mathbb{E}(X) = a, \quad var(X) = a, \quad mode(X) = a - 1 \text{ for } a \geq 1.$$

5.2.2 Noncentral Gamma distribution

Let $X \sim Ga(a, \lambda)$, i.e. let X be a random variable with a noncentral Gamma distribution with parameter $a > 0$ and with noncentrality parameter $\lambda \geq 0$.

- Distribution function of X:

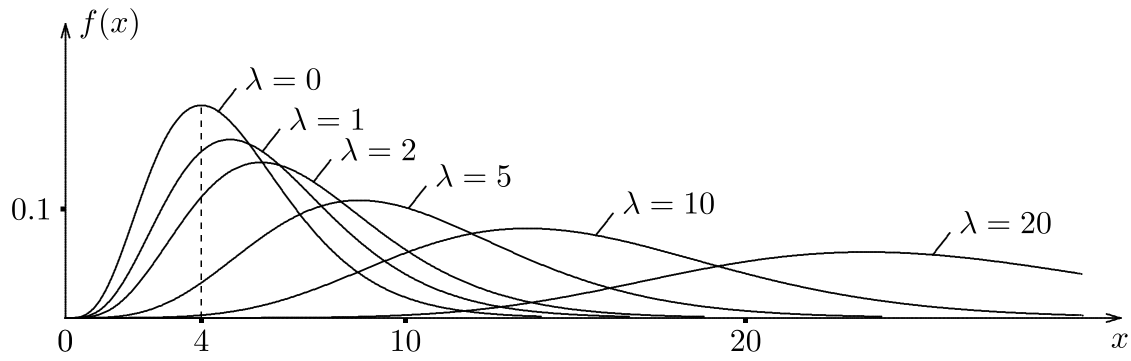
$$F(x|a, \lambda) = \mathbb{P}\{X < x\} = \sum_{j=0}^{\infty} p(j|\lambda) \cdot F(x|a + j)$$

where

$p(j|\lambda) = \dfrac{e^{-\lambda}\lambda^j}{j!}$ point probability of Poisson distribution $Po(\lambda)$;

$F(x|a)$ distribution function of central Gamma distribution with parameter a.

Fig. 6: Density of the noncentral Gamma distribution ($a = 5$)



- Expectation and variance:

$\mathbb{E}(X) = a + \lambda, \quad var(X) = a + 2\lambda.$

5.3 Chi-square Distribution

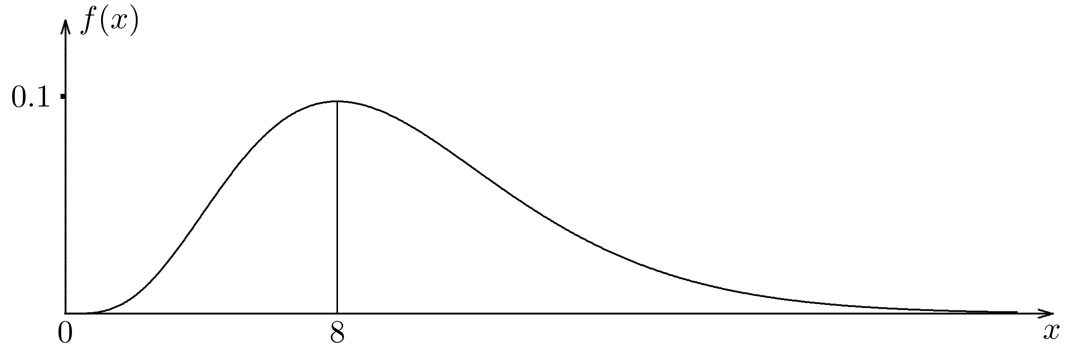
5.3.1 (Central) chi-square distribution

Let $X \sim \chi^2(n), \quad n = 1, 2, \ldots$, i.e. let X be a random variable with a χ^2 distribution with n degrees of freedom.

- Density function of X:

$$f(x) = \frac{1}{2\,\Gamma(n/2)} \left(\frac{x}{2}\right)^{(n-2)/2} e^{-x/2}, \quad x > 0.$$

Fig. 7: Density of the χ^2 distribution ($n = 10$)



- Constructive definition:

  Let

  $Z_1, \ldots, Z_n$ be independent random variables, each with $N(0, 1)$;

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

  Then:

  $X \sim \chi^2(n)$.

- Expectation, variance and mode:

  $\mathbb{E}(X) = n$;

  $var(X) = 2n$;

  $mode(X) = n - 2 \quad \text{for } n \geq 2$.

- Relation to the Gamma distribution:

  $X \sim \chi^2(n)$

  $\Rightarrow Y = X/2 \sim Ga(a)$ with $a = n/2$.

5.3.2 Noncentral chi-square distribution

Let $X \sim \chi^2(n,\lambda), \quad n = 1,2,\ldots, \quad \lambda \geq 0$,
i.e. let X be a random variable with a noncentral χ^2-distribution with n degrees of freedom and with noncentrality parameter λ.

- Distribution function of X:

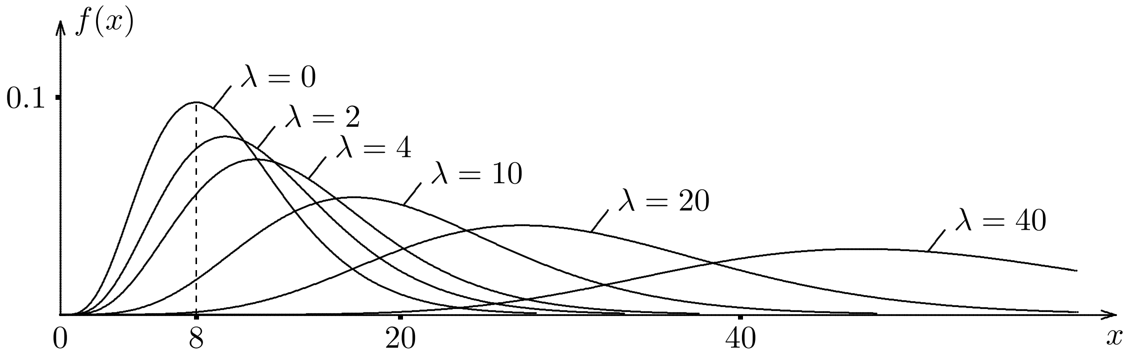
$$F(x|n,\lambda) = \mathbb{P}\{X < x\} = \sum_{j=0}^{\infty} p(j|\tilde{\lambda}) \cdot F(x|n+2j)$$

where $\tilde{\lambda} = \lambda/2$ and

$p(j|\lambda) = \dfrac{e^{-\lambda}\lambda^j}{j!}$ point probability of the Poisson distribution $Po(\lambda)$;

$F(x|n)$ distribution function of the central χ^2-distribution with n degrees of freedom.

Fig. 8: Density of the noncentral χ^2 distribution ($n = 10$)



- Constructive definition:

Let $Z_1,\ldots,Z_n$ be independent random variables where

$Z_j \sim N(\delta_j, 1), \quad j = 1,\ldots,n;$

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

Then: $\quad X \sim \chi^2(n,\lambda) \quad$ with $\quad \lambda = \sum_{j=1}^{n} \delta_j^2.$

- Expectation and variance:

$\mathbb{E}(X) = n + \lambda,$

$var(X) = 2(n + 2\lambda).$

- Relation to the noncentral Gamma distribution:

$X \sim \chi^2(n,\lambda)$

$\Rightarrow Y = X/2 \sim Ga(a,\tilde{\lambda}) \quad$ with $\quad a = n/2 \quad$ and $\quad \tilde{\lambda} = \lambda/2.$

5.4 Beta Distribution

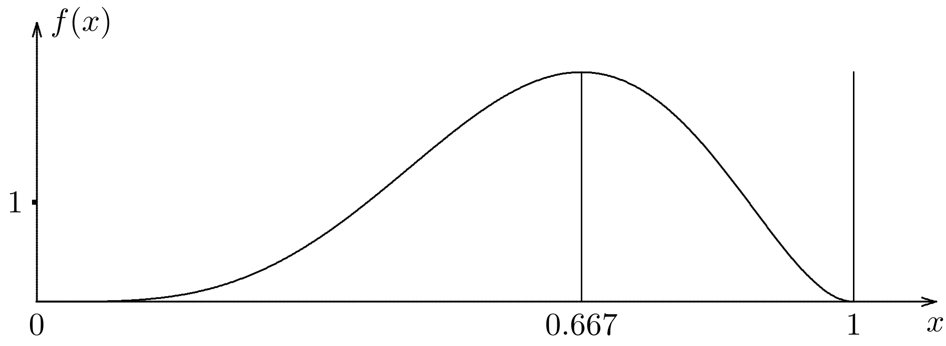
5.4.1 (Central) Beta distribution

Let $X \sim Be(a, b)$, $a, b > 0$, i.e. let X be a random variable with a Beta distribution with the parameters a and b.

- Density function of X:

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}, \quad 0 < x < 1.$$

Fig. 9: Density of the Beta distribution ($a = 5,\ b = 3$)



- Constructive definition:

  Let X_1, X_2 be independent random variables:

  $X_1 \sim Ga(a)$

  $X_2 \sim Ga(b)$.

  Then:

$$X = \frac{X_1}{X_1 + X_2} \sim Be(a, b).$$

- Expectation, variance and mode:

$$\mathbb{E}(X) = \frac{a}{a+b},$$

$$var(X) = \frac{a \cdot b}{(a+b)^2(a+b+1)},$$

$$mode(X) = \frac{a-1}{a+b-2} \quad \text{for } a \geq 1,\ \ b \geq 1,\ \ a+b > 2.$$

5.4.2 Noncentral Beta distribution

Let $X \sim Be(a, b, \lambda), \quad a, b > 0, \;\; \lambda \geq 0,$
i.e. let X be a random variable with a noncentral Beta distribution with parameters a and b and with noncentrality parameter λ.

- Distribution function of X:

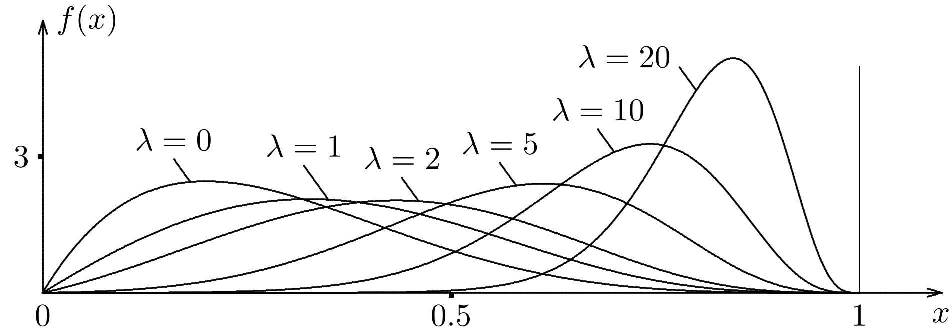
$$F(x|a,b,\lambda) = \mathbb{P}\{X < x\} = \sum_{j=0}^{\infty} p(j|\lambda) \cdot F(x|a+j, b)$$

where

$p(j|\lambda) = \dfrac{e^{-\lambda}\lambda^j}{j!}$ — point probability of Poisson distribution $Po(\lambda)$

$F(x|a,b)$ — distribution function of central Beta distribution with parameters a and b.

Fig. 10: Density of the noncentral Beta distribution ($a = 2, \; b = 5$)



- Constructive definition:

Let X_1, X_2 be independent random variables:

$X_1 \sim Ga(a, \lambda)$
$X_2 \sim Ga(b)$.

Then: $\quad X = \dfrac{X_1}{X_1 + X_2} \sim Be(a, b, \lambda).$

- Expectation and variance:

Exact simple formulas are not known to the author.
Approximations for large values of $a + \lambda$:

$$\mathbb{E}(X) \approx \frac{a+\lambda}{a+b+\lambda},$$

$$var(X) \approx \frac{(a+\lambda)\,b}{(a+b+\lambda)^3} + \frac{\lambda\, b^2}{(a+b+\lambda)^4}.$$

5.5 F Distribution

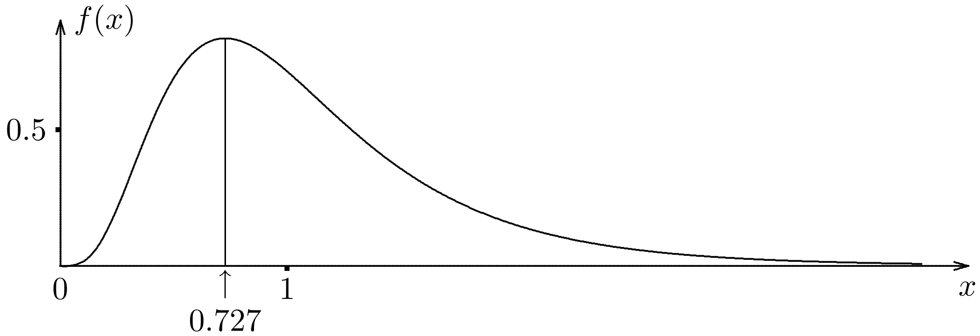
5.5.1 (Central) F distribution

Let $X \sim F(n_1, n_2), n_1, n_2 = 1, 2, \ldots,$, i.e. let X be a random variable with a F distribution with n_1 and n_2 degrees of freedom.

- Density function of X for $x > 0$:

$$f(x) = \frac{n_1}{n_2} \cdot \frac{\Gamma(\frac{n_1+n_2}{2})}{\Gamma(\frac{n_1}{2})\,\Gamma(\frac{n_2}{2})} \left(1 + \frac{n_1}{n_2}\,x\right)^{-(n_1+n_2)/2} \left(\frac{n_1}{n_2}\,x\right)^{(n_1-2)/2} .$$

Fig. 11: Density of the F distribution ($n_1 = 10,\ n_2 = 20$)



- Constructive definition:
  Let X_1, X_2 be independent random variables:
  $X_1 \sim \chi^2(n_1)$ and $X_2 \sim \chi^2(n_2)$.
  Then: $X = \dfrac{X_1/n_1}{X_2/n_2} \sim F(n_1, n_2)$.

- Expectation, variance and mode:

$$I\!E(X) = \frac{n_2}{n_2 - 2} \quad \text{for} \quad n_2 > 2,$$

$$var(X) = \frac{2(n_1 + n_2 - 2)}{n_1(n_2 - 4)} \left(\frac{n_2}{n_2 - 2}\right)^2 \quad \text{for} \quad n_2 > 4,$$

$$mode(X) = \frac{n_2(n_1 - 2)}{n_1(n_2 + 2)} \quad \text{for} \quad n_1 \geq 2.$$

- Relation to the Beta distribution:

$$X \sim F(n_1, n_2)$$

$$\Rightarrow Y = \frac{n_1 X}{n_2 + n_1 X} \sim Be(a, b) \quad \text{with} \quad a = n_1/2,\ b = n_2/2.$$

5.5.2 Noncentral F distribution

Let $X \sim F(n_1, n_2, \lambda)$, $n_1, n_2 = 1, 2, \ldots$, $\lambda \geq 0$, i.e. let X be a random variable with a noncentral F distribution with n_1 and n_2 degrees of freedom and with noncentrality parameter λ.

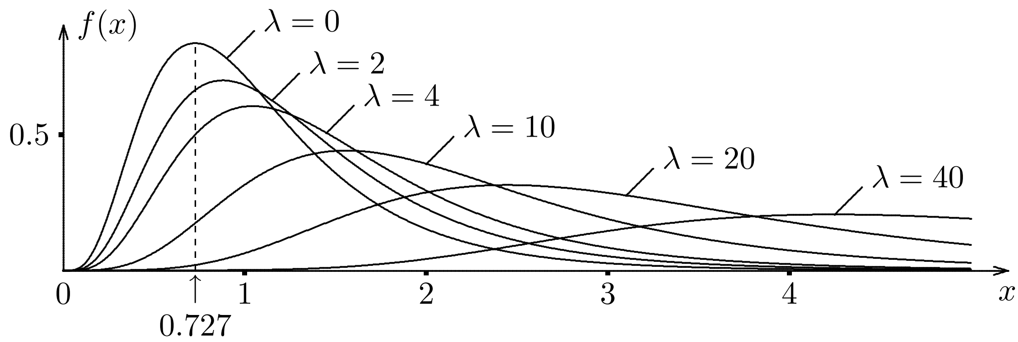
- Distribution function of X: (*)

$$F(x|n_1, n_2, \lambda) = I\!P\{X < x\} = \sum_{j=0}^{\infty} p(j|\tilde{\lambda}) \cdot F(\frac{n_1 x}{n_1 + 2j} \mid n_1 + 2j, n_2)$$

where $\tilde{\lambda} = \lambda/2$ and

$p(j\vert\lambda) = \dfrac{e^{-\lambda}\lambda^j}{j!}$	point probability of the Poisson distribution $Po(\lambda)$
$F(x\vert n_1, n_2)$	distribution function of the central F distribution with n_1 and n_2 degrees of freedom.

Fig. 12: Density of the noncentral F distribution ($n_1 = 10,\ n_2 = 20$)



- Constructive definition:

Let X_1, X_2 be independent random variables:

$X_1 \sim \chi^2(n_1, \lambda)$

$X_2 \sim \chi^2(n_2)$.

Then:

$$X = \frac{X_1/n_1}{X_2/n_2} \sim F(n_1, n_2, \lambda).$$

(*) Formula 26.6.18 in Abramowitz-Stegun(1972) reads (in our notation):
$F(x|n_1, n_2, \lambda) = \sum_{j=0}^{\infty} p(j|\lambda) \cdot F(x|n_1 + 2j, n_2)$.

Bablok(1988) has observed, that this relation is not correct.

- Expectation and variance:
$\mathbb{E}(X) = \frac{n_2}{n_2 - 2}(1 + \rho) \quad \text{with } \rho = \lambda/n_1, \quad \text{for} \quad n_2 > 2$
$$var(X) = \frac{2(n_1 + n_2 - 2)}{n_1(n_2 - 4)} \left(\frac{n_2}{n_2 - 2}\right)^2 \left(1 + 2\rho + \frac{n_1}{n_1 + n_2 - 2}\rho^2\right)$$
for $n_2 > 4$.

- Relation to the noncentral Beta distribution:
$X \sim F(n_1, n_2, \lambda)$
$\Rightarrow Y = \frac{n_1 X}{n_2 + n_1 X} \sim Be(a, b, \tilde{\lambda})$ with $a = n_1/2,\ b = n_2/2,\ \tilde{\lambda} = \lambda/2.$

5.6 t Distribution (Student's Distribution)

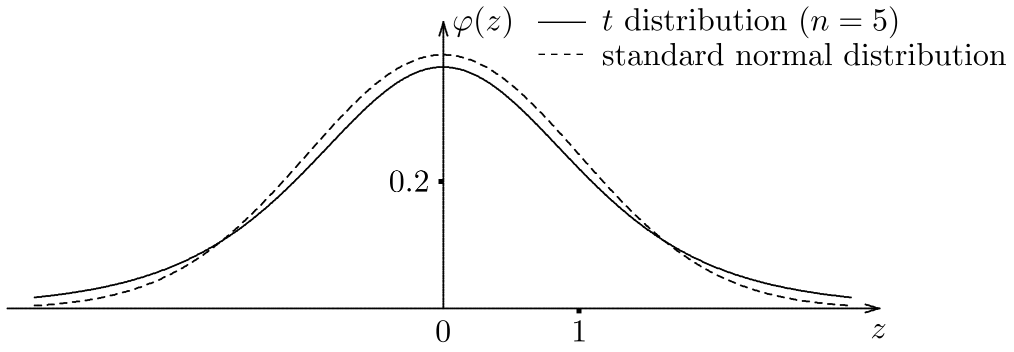
5.6.1 (Central) t distribution

Let $X \sim t(n),\ n = 1, 2, \ldots$, i.e. let X be a random variable with a t distribution with n degrees of freedom.

- Density function of X:

$$f(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})}\left(1 + \frac{x^2}{n}\right)^{-(n+1)/2}, \quad -\infty < x < \infty.$$

Fig. 13: Density of the t distribution (Student's distribution)



- Constructive definition:

  Let Y and Z be independent random variables:

  $Z \sim N(0,1)$ and $Y \sim \chi^2(n)$.

  Then: $X = \dfrac{Z}{\sqrt{Y/n}} \sim t(n)$.

- Expectation and variance:

  $E(X) = 0$ for $n > 1$,

  $var(X) = \dfrac{n}{n-2}$ for $n > 2$.

- Relation to the F distribution:

  $X \sim t(n)$

  $\Rightarrow Y = X^2 \sim F(1, n)$.

- Relation to the Beta distribution:

  $X \sim t(n)$

  $\Rightarrow Y = \dfrac{X^2}{X^2 + n} \sim Be(a, b)$ with $a = 1/2,\ b = n/2$.

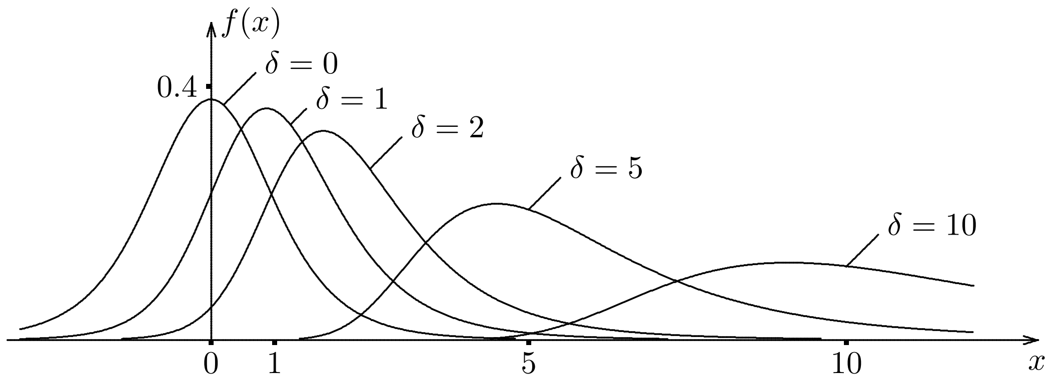
5.6.2 Noncentral t distribution

Let $X \sim t(n,\delta)$, $n = 1, 2, \ldots$, $\delta \in I\!R$, i.e. let X be a random variable with a noncentral t distribution with n degrees of freedom and with noncentrality parameter δ.

- Distribution function of X for $-\infty < x < \infty$: (*)

$$F(x|n,\delta) = I\!P\{X < x\} = \frac{\sqrt{2\pi}}{2^{(n-2)/2}\,\Gamma(\frac{n}{2})} \int_0^\infty s^{n-1}\,\Phi(\frac{sx}{\sqrt{n}} - \delta)\,\varphi(s)\,ds.$$

Fig. 14: Density of the noncentral t distribution ($n = 5$)



- Constructive definition:

  Let Y and Z be independent random variables:

  $Z \sim N(\delta, 1)$

  $Y \sim \chi^2(n)$.

  Then: $X = \dfrac{Z}{\sqrt{Y/n}} \sim t(n,\delta)$.

- Expectation and variance:

$$I\!E(X) = \sqrt{\tfrac{n}{2}}\,\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\,\delta \quad \text{for} \quad n > 1;$$

(*) Bablok(1988) has observed that the second equation of 26.7.9 in Abramowitz-Stegun(1972) computes the two-sided probabilities $I\!P\{|X| < x\}$. Due to the asymmetry of the noncentral t distribution the one-sided probabilities cannot be derived from this equation.

$$var(X) = \frac{n}{n-2} + \delta^2 \left[\frac{n}{n-2} - \frac{n}{2} \left(\frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \right)^2 \right] \quad \text{for} \quad n > 2.$$

- Relation to the noncentral F distribution:

  $X \sim t(n, \delta)$

  $\Rightarrow Y = X^2 \sim F(1, n, \lambda) \quad \text{with } \lambda = \delta^2.$

- Relation to the noncentral Beta distribution:

  $X \sim t(n, \delta)$

  $$\Rightarrow Y = \frac{X^2}{X^2 + n} \sim Be(a, b, \lambda) \text{ with } \quad a = 1/2, \ \ b = n/2, \ \ \lambda = \delta^2/2.$$

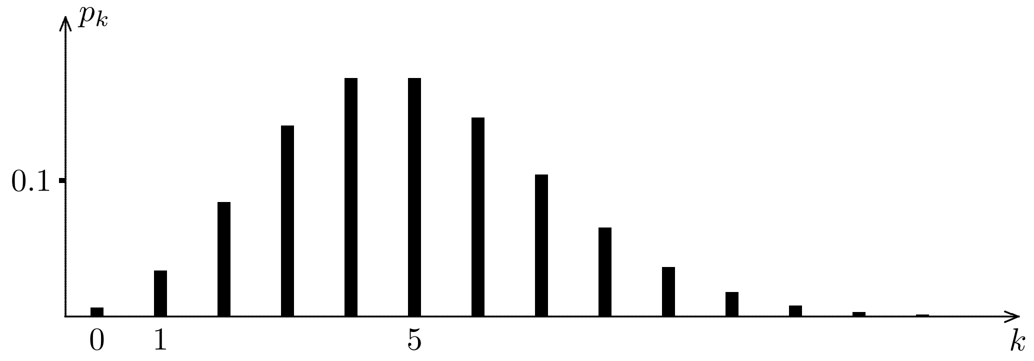
5.7 Poisson Distribution

Let $X \sim Po(\lambda)$, $\lambda > 0$, i.e. let X be a random variable with a Poisson distribution with parameter λ.

- Point probability of X:

$$p_k = I\!P\{X = k\} = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, \ldots .$$

Fig. 15: Poisson distribution ($\lambda = 5$)



- Expectation, variance and mode:

$I\!E(X) = \lambda, \quad var(X) = \lambda,$

$mode(X) = \lfloor \lambda \rfloor.$ (*)

- Relation to the Gamma distribution:

Let

$X \sim Po(\lambda).$

Then we have for $k = 1, 2, \ldots$:

$I\!P\{X \geq k\} = I\!P\{Y < \lambda\}$ where $Y \sim Ga(k).$

- Relation to the chi-square distribution:

Let

$X \sim Po(\lambda).$

Then we have for $k = 1, 2, \ldots$:

$I\!P\{X \geq k\} = I\!P\{Y < 2\lambda\}$ where $Y \sim \chi^2(2k).$

(*) $\lfloor \lambda \rfloor$ = integer part of λ. If λ is an integer, there are two modes: λ and $\lambda - 1$.

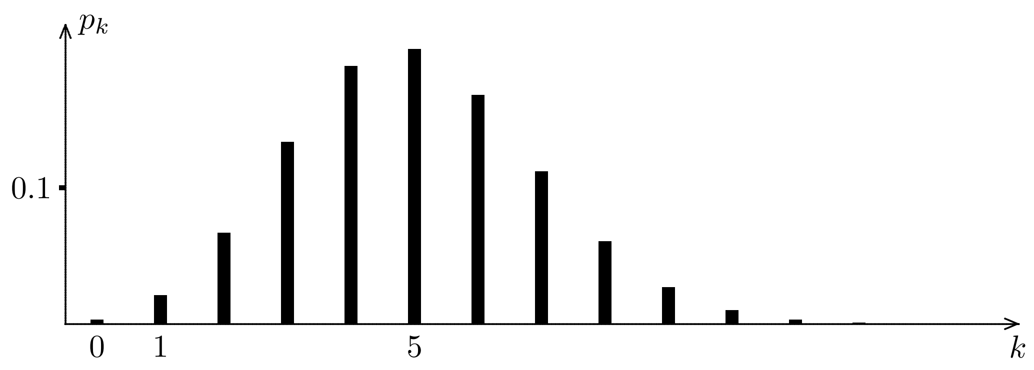
5.8 Binomial Distribution

Let $X \sim Bi(n, \vartheta),\ n = 1, 2, \ldots,\ 0 < \vartheta < 1$, i.e. let X be a random variable with a binomial distribution with the parameters n and ϑ.

- Point probability of X:

$$p_k = \mathbb{P}\{X = k\} = \binom{n}{k} \vartheta^k (1-\vartheta)^{n-k}, \quad k = 0, 1, \ldots, n.$$

Fig. 16: Binomial distribution ($n = 20$, $\theta = 0.25$)



- Constructive definition:

  We consider a random experiment with n independent trials; in each trial a certain random event A can occur (the urn model with replacement is a special case of such an experiment). Let

$\vartheta = \mathbb{P}(A) =$	probability of A in a single trial;
$n =$	total number of trials;
$X=$	number of successes
	(= number of trials where A occurs).

  Then: $X \sim Bi(n, \vartheta)$.

- Expectation, variance and mode:

  $\mathbb{E}(X) = n\vartheta$;

  $var(X) = n\vartheta(1-\vartheta)$;

  $mode(X) = \lfloor y \rfloor$, where $y = (n+1)\vartheta$. (*)

(*) $\lfloor y \rfloor$ = integer part of y. If y is an integer, there are two modes: y and $y - 1$.

- Relation to the Beta distribution:
  Let
  $$X \sim Bi(n, \vartheta).$$
  Then we have for $k = 1, \ldots, n$:
  $$\mathbb{P}\{X \geq k\} = \mathbb{P}\{Y < \vartheta\}$$
  where
  $$Y \sim Be(a, b) \quad \text{with} \quad a = k \quad \text{and} \quad b = n - k + 1.$$

- Relation to the F distribution:
  Let
  $$X \sim Bi(n, \vartheta).$$
  Then we have for $k = 1, \ldots, n$:
  $$\mathbb{P}\{X \geq k\} = \mathbb{P}\{Y < \frac{n-k+1}{k} \cdot \frac{\vartheta}{1-\vartheta}\}$$
  where
  $$Y \sim F(n_1, n_2) \quad \text{with} \quad n_1 = 2k \quad \text{and} \quad n_2 = 2(n - k + 1).$$

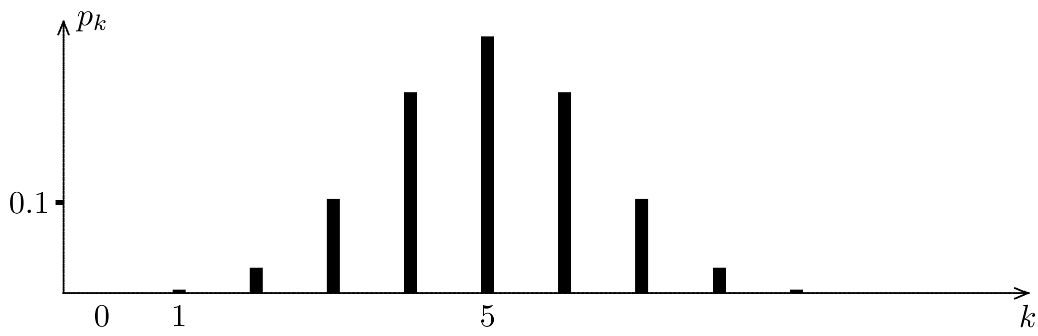
5.9 Hypergeometric Distribution

Let $X \sim H(N, M, n)$, $N = 2, 3, \ldots;\ 0 < n, M < N$,
i.e. let X be a random variable with a hypergeometric distribution with the parameters N, M, n.

- Point probability of X:

$$p_k = \mathbb{P}\{X = k\} = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, \quad 0 \le k \le M, \quad 0 \le n-k \le N-M.$$

Fig. 17: Hypergeometric distribution with ($N = 40$, $M = 10$, $n = 20$)



- Constructive definition:

  From an urn with white and black balls a random sample is drawn without replacement, and let

  N = total number of balls in the urn;
  M = number of white balls in the urn;
  n = sample size (number of balls drawn without replacement);
  X = number of white balls in the sample.

  Then: $X \sim H(N, M, n)$.

- Expectation, variance and mode:

$$\mathbb{E}(X) = n\vartheta, \quad \text{with } \vartheta = \frac{M}{N};$$

$$var(X) = n\vartheta(1-\vartheta)\,\frac{N-n}{N-1};$$

$$mode(X) = \lfloor y \rfloor, \quad \text{where } y = \frac{(M+1)(n+1)}{N+2}. \quad ^{(*)}$$

(*) $\lfloor y \rfloor$ = integer part of y. If y is an integer, there are two modes: y and $y-1$.

6 Basic Ideas of the Computations

The numerical computation of all our nine distributions relies upon the same basic ideas. Here we want to show up these ideas with the Poisson distribution.

6.1 Aim of the Computations

Let X be a random variable with a Poisson distribution with parameter $\lambda > 0$. Then

$$p_k = \mathbb{P}\{X = k\} = \frac{e^{-\lambda}\lambda^k}{k!}, \qquad k = 0, 1, \dots$$

For given λ and k ($\lambda, k \leq 10^8$, about) we want to compute the three probabilities:

a) $\mathbb{P}\{X < k\}$ = lower tail probability
b) $\mathbb{P}\{X > k\}$ = upper tail probability
c) $\mathbb{P}\{X = k\}$ = point probability.

Each of the three probabilities should be computed with a *relative error* $\leq \varepsilon$ (e.g. $\varepsilon = 10^{-6}$, i.e. with six correct significant digits); this should be true as long as the correct probability is larger than a minimum value $r_{\min}$ (with ELV $r_{\min} = 10^{-100}$); probabilities smaller than $r_{\min}$ are set to zero. We want to point out that only the smaller of the two probabilities a) and b) has to be computed directly; the other one can then be found by complementation as all three probabilities sum up to one.

6.2 Naive Method

The following recurrence relation holds true:

$$\begin{aligned} p_k &= \frac{\lambda}{k}\, p_{k-1}, \qquad k = 1, 2, \dots \\ p_0 &= e^{-\lambda}. \end{aligned} \tag{6.1}$$

Thus we find

by summation: $\mathbb{P}\{X < k\} = p_0 + p_1 + \cdots + p_{k-1}$
by subtraction: $\mathbb{P}\{X > k\} = 1 - \mathbb{P}\{X \leq k\}$,

and our numerical problem seems to be solved. The following remarks show that this is not the case:

a) We have

$$p_0 = e^{-\lambda} < 10^{-100} \quad \text{if} \quad \lambda > 230.26$$
$$< 10^{-1000} \quad \text{if} \quad \lambda > 2302.6$$

and this shows, that for large λ we face underflow problems when computing p_0, as real numbers that are too small are rounded down to zero by the computer (underflow).

b) For given λ we have $\mathbb{P}\{X \le k\} \to 1$ for $k \to \infty$. So the formula $\mathbb{P}\{X > k\} = 1 - \mathbb{P}\{X \le k\}$ suffers from numerical cancellation for large k, and upper tail probabilities cannot be computed with a given *relative* error in this way if these probabilities are too small.

c) Even if the naive method works, it is inefficient in most cases, as the correct result can be found with much fewer steps (see Table 2 below).

6.3 Approximation of the Poisson by the Normal Distribution

One could think, that the exact computation of the Poisson distribution for large values of λ ($\lambda \ge 100$ or so) is not necessary at all, as the Poisson distribution can be approximated by the normal distribution due to the central limit theorem:

$$\mathbb{P}\{X \le k\} \approx \Phi(z_k) \tag{6.2}$$

where

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt \qquad \text{and} \qquad z_k = \frac{k + \frac{1}{2} - \lambda}{\sqrt{\lambda}}$$

(Φ = distribution function of the standard normal distribution). Table 1a – 1c informs about the relative error of this approximation, where this error is defined as

$$rel.err. = \frac{app - exa}{exa} \quad \text{with} \quad exa = \mathbb{P}\{X \le k\} \quad \text{and} \quad app = \Phi(z_k).$$

We see, that even for $\lambda = 10^6$ our aim (absolute value of *relative error* $\le \varepsilon = 10^{-6}$) cannot be achieved with this simple approximation, not even in the central part of the distribution. For $\lambda = k = 10^6$ the relative error is still 134ε; as the error behaves like $1/\sqrt{\lambda}$ for large λ, we ought to have $\lambda \ge (134)^2 \cdot 10^6 \approx 18 \cdot 10^9$ for the relative error to be $\le \varepsilon = 10^{-6}$ at least for $\lambda = k$.

Tabelle 1a: Normal approximation (6.2) for $\lambda = 100$

k	z_k	$\mathbb{P}\{X \leq k\}$	$\Phi(z_k)$	*rel.err.*
100	0.05	0.526 562	0.519 939	−0.0126
95	−0.45	0.331 192	0.326 355	−0.0146
90	−0.95	0.171 385	0.171 056	−0.0019
85	−1.45	0.707 505 E−1	0.735 293 E−1	0.0393
80	−1.95	0.226 492 E−1	0.255 881 E−1	0.130
75	−2.45	0.547 266 E−2	0.714 281 E−2	0.305
70	−2.95	0.971 444 E−3	0.158 887 E−2	0.636
60	−3.95	0.108 122 E−4	0.390 756 E−4	2.61
50	−4.95	0.240 159 E−7	0.371 067 E−6	14.45

Tabelle 1b: Normal approximation (6.2) for $\lambda = 10^4$

k	z_k	$\mathbb{P}\{X \leq k\}$	$\Phi(z_k)$	*rel.err.*
10 000	0.005	0.502 660	0.501 995	−0.001 32
9 950	−0.495	0.310 744	0.310 300	−0.001 43
9 900	−0.995	0.159 871	0.159 868	−0.000 02
9 850	−1.495	0.671 865 E−1	0.674 572 E−1	0.004 03
9 800	−1.995	0.227 492 E−1	0.230 214 E−1	0.0120
9 750	−2.495	0.614 319 E−2	0.629 786 E−2	0.0252
9 700	−2.995	0.131 286 E−2	0.137 222 E−2	0.0452
9 600	−3.995	0.290 528 E−4	0.323 471 E−4	0.113
9 500	−4.995	0.237 938 E−6	0.294 179 E−6	0.236

Tabelle 1c: Normal approximation (6.2) for $\lambda = 10^6$

k	z_k	$\mathbb{P}\{X \leq k\}$	$\Phi(z_k)$	*rel.err.*
1 000 000	0.0005	0.500 266	0.500 199	−0.000 134
999 500	−0.4995	0.308 785	0.308 714	−0.000 230
999 000	−0.9995	0.158 776	0.158 776	0
998 500	−1.4995	0.668 450 E−1	0.668 720 E−1	0.000 404
998 000	−1.9995	0.227 501 E−1	0.227 771 E−1	0.001 19
997 500	−2.4995	0.620 308 E−2	0.621 843 E−2	0.002 47
997 000	−2.9995	0.134 620 E−2	0.135 212 E−2	0.004 40
996 000	−3.9995	0.314 042 E−4	0.317 382 E−4	0.0106
995 000	−4.9995	0.281 482 E−6	0.287 396 E−6	0.0210

6.4 Computation of Upper Tail Probabilities

We want to show how to compute upper tail probabilities reliably. From the recurrence relation (6.1) we have

$$\begin{aligned} &p_j = c_j \cdot p_{j-1} \quad \text{with} \quad c_j = \lambda/j \quad \text{for} \quad j = 1, 2, \ldots \\ &c_j \downarrow 0 \quad \text{for} \quad j \uparrow \infty \quad (\text{monotony of } c_1, c_2, \cdots). \end{aligned} \tag{6.3}$$

So the upper tail probability becomes

$$\begin{aligned} I\!\!P\{X > k\} &= p_{k+1} + p_{k+2} + p_{k+3} + \cdots \\ &= p_k \cdot c_{k+1} + p_k \cdot c_{k+1}c_{k+2} + p_k \cdot c_{k+1}c_{k+2}c_{k+3} + \cdots \\ &= p_k \cdot I_k \end{aligned}$$

where

$$\begin{aligned} p_k &= \frac{e^{-\lambda}\lambda^k}{k!} \\ I_k &= c_{k+1} + c_{k+1}c_{k+2} + c_{k+1}c_{k+2}c_{k+3} + \cdots \end{aligned}$$

The computation of p_k seems to be obvious, but we will see in section 6.6 that this is not the case if λ and k are large. Here we want to deal with the computation of I_k. I_k looks like a geometric series in c (namely: $c + c^2 + c^3 + \cdots$), and due to the monotony of the sequence $c_1, c_2, \cdots$ the series I_k is convergent for any $\lambda > 0$ and $k \geq 1$. From the monotony of the sequence c_j and the definition of I_k we obviously have

$$\begin{aligned} &(i) \quad I_k \downarrow 0 \quad \text{for} \quad k \uparrow \infty \\ &(ii) \quad I_{k-1} = c_k(1 + I_k) \quad \text{for} \quad k = 1, 2, \ldots \end{aligned}$$

Now we consider the following algorithm for computing an approximation to I_k for given λ and k:

$$\begin{aligned} &\text{Set} \quad \tilde{I}_{k+n} = 0 \quad \text{for some positive integer } n. \\ &\text{Compute for } j = k+n, k+n-1, \ldots, k+1: \\ &\qquad \tilde{I}_{j-1} = c_j(1 + \tilde{I}_j) \quad \text{where} \quad c_j = \lambda/j. \\ &\text{Then } \tilde{I}_k \text{ is an approximation to } I_k. \end{aligned} \tag{6.4}$$

This strange algorithm starts somewhere at $k + n$ with a starting value $\tilde{I}_{k+n} = 0$, that is obviously wrong as $I_k > 0$ for any $k \geq 1$, and it proceeds backwards in n steps to the final value $\tilde{I}_k$. We call this algorithm *open backward recursion* with n steps, and we have to think about the approximation error of this algorithm. We obviously have

$$I_k = c_{k+1}(1 + I_{k+1})$$
$$\tilde{I}_k = c_{k+1}(1 + \tilde{I}_{k+1}),$$

and from this we derive

$$I_k - \tilde{I}_k = c_{k+1}(I_{k+1} - \tilde{I}_{k+1}).$$

By repetition we find

$$I_k - \tilde{I}_k = c_{k+1}c_{k+2}\cdots c_{k+n}(I_{k+n} - \tilde{I}_{k+n}).$$

Now, $\tilde{I}_{k+n} = 0$, and so we have

$$\frac{I_k - \tilde{I}_k}{I_k} = c_{k+1}c_{k+2}\cdots c_{k+n}\frac{I_{k+n}}{I_k}.$$

Due to the monotony of the sequence I_k we have $I_{k+n} < I_k$ and so we obtain

$$0 < \frac{I_k - \tilde{I}_k}{I_k} < prod,$$

where

$$prod = c_{k+1}c_{k+2}\cdots c_{k+n}.$$

So we have found the following algorithm for computing I_k for given λ, k $(\lambda > 0, k \geq 1)$ and for given ε (e.g. $\varepsilon = 10^{-6}$):

1. Determine n such that

   $$prod = c_{k+1}c_{k+2}\cdots c_{k+n} < \varepsilon.$$

2. Determine the approximation $\tilde{I}_k$ by the open backward recursion (6.4) with n steps. (6.5)

Then: $$0 < \frac{I_k - \tilde{I}_k}{I_k} < \varepsilon.$$

Thus $\tilde{I}_k$ is an approximation to I_k with a *relative* error $< \varepsilon$. Note that the algorithm (6.5) is particularly efficient if $k \geq \lambda$ as in this case all coefficients $c_{k+1}, \ldots, c_{k+n}$ are smaller than 1. And it is just the case $k \geq \lambda$ where we need this algorithm as in this case the upper tail probabilities are smaller than the lower ones. For given λ the necessary number of steps to achieve a relative accuracy of at least ε becomes the smaller the larger k i.e. the algorithm becomes more efficient for extreme values of k. On condition that $k \geq \lambda$ the worst case as concerns efficiency is met for $k = \lambda$ (if λ is an integer). Table 2 gives the necessary number of steps in this worst case for different values of λ and ε.

Table 2: Necessary number of steps in the worst case $k = \lambda$:
(forw.) = forward recusion (6.5)
(backw.) = backward recursion (6.7)

	$\varepsilon = 10^{-3}$		$\varepsilon = 10^{-6}$		$\varepsilon = 10^{-9}$	
$\lambda\ (= k)$	forw.	backw.	forw.	backw.	forw.	backw.
100	36	39	49	57	58	71
10 000	370	374	522	530	638	651
1 000 000	3716	3719	5253	5261	6432	6445
$\lambda \to \infty$	$3.717\sqrt{\lambda}$		$5.257\sqrt{\lambda}$		$6.438\sqrt{\lambda}$	

The necessary number of steps for $\lambda \to \infty$ is $y\sqrt{\lambda}$,
where $\varepsilon = e^{-y^2/2}$ (cf. Knüsel(1986)).

6.5 Computation of Lower Tail Probabilities

Now we want to deal with the computation of the lower tail probabilities $I\!P\{X < k\}$. From the recurrence relation (6.1) we have

$$p_{j-1} = d_j \cdot p_j \quad \text{with} \quad d_j = j/\lambda \quad \text{for} \quad j = 1, 2, \ldots$$
$$d_j \downarrow 0 \quad \text{for} \quad j \downarrow 0 \quad (\text{monotony of } d_0, d_1, \cdots).$$

Therefore

$$\begin{aligned} I\!P\{X < k\} &= p_{k-1} + p_{k-2} + \cdots + p_0 \\ &= p_k \cdot d_k + p_k \cdot d_k d_{k-1} + \cdots + p_k \cdot d_k d_{k-1} \cdots d_1 \\ &= p_k \cdot J_k \end{aligned}$$

where

$$p_k = \frac{e^{-\lambda}\lambda^k}{k!}$$
$$J_k = d_k + d_k d_{k-1} + \cdots + d_k d_{k-1} \cdots d_1.$$

The relation $I\!P\{X < k\} = p_k J_k$ is also true for $k = 0$ if we set $J_0 = 0$. From the monotony of the sequence d_j and the definition of J_k we obviously have

$$(i) \quad J_k \downarrow 0 \quad \text{for} \quad k \downarrow 0$$
$$(ii) \quad J_k = d_k(1 + J_{k-1}) \quad \text{for} \quad k = 1, 2, \ldots \quad (J_0 = 0).$$

We consider the following algorithm for computing an approximation to J_k for given λ and k (*open forward recursion*):

$$\begin{aligned}&\text{Set } \tilde{J}_{k-n} = 0 \text{ for some positive integer n.}\\ &\text{Compute for } j = k-n+1, k-n+2, \ldots, k\text{:}\\ &\qquad \tilde{J}_j = d_j(1+\tilde{J}_{j-1}) \quad \text{where} \quad d_j = j/\lambda.\\ &\text{Then } \tilde{J}_k \text{ is an approximation to } J_k.\end{aligned} \tag{6.6}$$

We obviously have

$$J_k = d_k(1 + J_{k-1})$$
$$\tilde{J}_k = d_k(1 + \tilde{J}_{k-1}).$$

and from this we derive iteratively

$$\begin{aligned}J_k - \tilde{J}_k &= d_k(J_{k-1} - \tilde{J}_{k-1})\\ &= d_k d_{k-1} \cdots d_{k-n+1}(J_{k-n} - \tilde{J}_{k-n}).\end{aligned}$$

Now, $\tilde{J}_{k-n} = 0$, and therefore

$$\frac{J_k - \tilde{J}_k}{J_k} = d_k d_{k-1} \cdots d_{k-n+1} \frac{J_{k-n}}{J_k}.$$

Due to the monotony of J_k we have $J_{k-n} < J_k$ and so we obtain

$$0 < \frac{J_k - \tilde{J}_k}{J_k} < prod$$

where

$$prod = d_k d_{k-1} \cdots d_{k-n+1}.$$

So we have found the following algorithm to compute J_k for given λ and k ($\lambda > 0, k \geq 1$) and for given ε (e.g. $\varepsilon = 10^{-6}$):

1. Determine n such that

   $prod = d_k d_{k-1} \cdots d_{k-n+1} < \varepsilon.$
2. Determine the approximation $\tilde{J}_k$ by the open forward recursion (6.6) with n steps. (6.7)

Then: $0 < \dfrac{J_k - \tilde{J}_k}{J_k} < \varepsilon.$

Thus $\tilde{J}_k$ is an approximation to J_k with a relative error $< \varepsilon$. Note that the required number n of steps is $\leq k$; for $n = k$ the algorithm starts with the correct starting value $\tilde{J}_0 = J_0 = 0$ and so the exact value of J_k (up to numerical rounding errors) is computed by complete forward recursion. In this case the algorithm is equivalent to the naive method described in section 6.2. Also note that the algorithm (6.7) is particularly efficient for $k \leq \lambda$ as in this case all coefficients $d_k, d_{k-1}, \ldots$ are ≤ 1. And it is just the case $k \leq \lambda$ where we need the algorithm as the lower tail probabilities are smaller than the upper ones. The necessary numer of steps in the worst case $k = \lambda$ for different values of λ and ε is also given in Table 2.

6.6 Computation of Point Probabilities

The computation of the point probabilities

$$p_k = \mathbb{P}\{X = k\} = \frac{e^{-\lambda}\lambda^k}{k!}$$

seems to be obvious. In order to avoid underflow problems for large λ and k one first computes the logarithm

$$\ln(p_k) = k \cdot \ln(\lambda) - \ln(k!) - \lambda,$$

and then the result is found by exponentiation: $p_k = e^{\ln(p_k)}$. For $k = \lambda = 10^6$ e.g. we obtain

$$\begin{aligned} k \cdot \ln(\lambda) &= 13\,816\,510.6 \\ \ln(k!) + \lambda &= 13\,816\,518.4 \\ \ln(p_k) &= \text{-7.8.} \end{aligned}$$

This example shows that severe numeric cancellation can occur with this method of computing p_k. Note that the number of correct digits after the decimal point in $\ln(p_k)$ determines the number of correct significant digits of p_k, as the *absolute* error of $\ln(p_k)$ is equal to the *relative* error of p_k. So, if we want to compute p_k with at least six correct significant digits in our example, we have to compute $\ln(\lambda)$ and $\ln(k!)$ with at least 14 correct significant digits. As it is not easy to extend the computer arithmetic beyond the standard precision (about 15 significant digits) one has to think about other measures to reduce cancellation (cf. Knüsel, (1986)).

6.7 Analogous Methods for the Other Distributions

Quite the same way as with the Poisson distribution we can compute the binomial and hypergeometric distribution, as similar recurrence relations hold true. The relation between the Poisson and Gamma distribution on one side, and between the binomial and Beta distribution on the other side, make it plausible that similar algorithms also work with the Gamma and Beta distribution, at least as long as the parameters a and b are integers. For real parameters one has to compute reliable starting values if the parameters are too small for the open recursions to give the required relative accuracy. The χ^2, t and F distributions can be derived from the Gamma and Beta distribution and offer no new problems. The standard normal distribution can be derived from the χ^2 distribution with one degree of freedom, and the noncentral distributions can be computed by means of the corresponding central distributions as we have seen in section 4.

7 References

Abramowitz, M., Stegun, I.A. (1972). *Handbook of Mathematical Functions*. Ninth Printing, Dover Publications, New York.

Bablok, B. (1988). *Numerische Berechnung nichtzentraler statistischer Verteilungen*. Diploma Thesis at the Department of Statistics, Univeristy of Munich, Germany.

Griffith, P. and Hill, I.D. (1985). *Applied Statistics Algorithms*. The Royal Statistical Society, London.

Knüsel, L. (1986). Computation of the Chi-square and Poisson Distribution. *SIAM Journal on Scientific and Statistical Computing* **7**, 1022–1036.

Knüsel, L. and Bablok, B. (1996). Computation of the Noncentral Gamma Distribution. *SIAM Journal on Scientific Computing* **17**, 1224–1231.