Mathematical Formulas and Other Valuable Knowledge

that I have found Useful for Myself
and Decided to Write Down

Ott Toomet

Seattle, October 27, 2017

Contents

1 Geometry

1.1 Koordinaatteisendused

1.1.1 Pinnaelement koordinaatteisendusel

Üleminekul koordinaatsüsteemilt (x, y) uutele koordinaatidele $(u = u(x, y), v = v(x, y))$ teisenevad väikesed koordinaatide sihilised lõigud nii:

$$\begin{pmatrix} \mathrm{d}u \\ \mathrm{d}v \end{pmatrix} = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} \begin{pmatrix} \mathrm{d}x \\ \mathrm{d}y \end{pmatrix}, \tag{1.1.1}$$

kus paremal poolel esimene on vastavate osatuletiste maatriks (jakobjaan $\mathcal{J}$). Pinnaelement avaldub

$$\mathrm{d}S = \mathrm{d}x\mathrm{d}y = \frac{\mathrm{d}u\mathrm{d}v}{|\det \mathcal{J}|}, \tag{1.1.2}$$

kus $|\det \mathcal{J}| = \mathrm{d}u\mathrm{d}v \sin(\widehat{u, v})$ on kahe lühikese lõigu du ja dv poolt defineeritud elementaarpindala. Analoogiline seos kehtib ka kõrgemate mõõtmete puhul.

1.2 Kolmnurk

1.2.1 Seos kolmnurga külgede ja nurkade vahel

$$a^2 + b^2 - 2ab \cos \gamma = c^2. \tag{1.2.1}$$

Täisnurksel kolmnurgal, kui $\gamma = 90°$, kehtib *Pythagorase teoreem*:

$$a^2 + b^2 = c^2. \tag{1.2.2}$$

1.2.2 Kolmnurga pindala

Kui kolmnurk on tasandil antud kolme punktiga $(0, 0)$, (X_A, Y_A) ning (X_B, Y_B) siis kolmnurga pindala on

$$S = \frac{1}{2} |X_A \cdot Y_B - Y_A \cdot X_B| = \frac{1}{2} \mathrm{abs} \begin{vmatrix} X_A & Y_A \\ X_B & X_B \end{vmatrix}. \tag{1.2.3}$$

Tõestus: joonista kolmnurk välja ja vaata, millised pindalad kirjeldab determinant.

1.3 Muud

1.3.1 Joone kõverus

Parameetriliselt antud joone kõverus:

$$k = \frac{x'y'' - y'x''}{(x'^2 + y'^2)^{3/2}} \tag{1.3.1}$$

2 Functions

2.1 Algebraic Functions

Gamma Function $\Gamma(\cdot)$

$$\Gamma(x) = \int_0^\infty t^{x-1} \mathrm{e}^{-1} \, \mathrm{d}t \tag{2.1.1}$$

Properties:

$$\Gamma(n) = (n-1)! \tag{2.1.2}$$

$$\Gamma(n+1) = n\Gamma(n) \tag{2.1.3}$$

Modified Bessel Function of the First Kind $I_\alpha(\cdot)$

$$I_\alpha(x) = \sum_{k=0}^{\infty} \frac{1}{k!(\alpha+k)!} \left(\frac{x}{2}\right)^{2k+\alpha} \tag{2.1.4}$$

3 Calculus

3.1 Logarithm

Properties

$$\log x^\alpha = \alpha \log x \qquad \text{and} \qquad \log(xy) = \log x + \log y \tag{3.1.1}$$

3.2 Limits

$$\lim_{x\to 0} x \log x = 0 \tag{3.2.1}$$

Proof. Write $x \log x = (\log x)/(1/x)$ and use L'Hospital's rule. □

$$\lim_{x\to\infty} \frac{1}{x}\frac{\phi(x)}{1-\Phi(x)} = 1, \tag{3.2.2}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are normal density and distribution functions.

Proof. The fraction can be written as

$$\frac{\phi(x)}{1-\Phi(x)} = \frac{\phi(x)}{\int_x^\infty \phi(s)\mathrm{d}s}.$$

The integral can be expressed as the Rieman limit

$$\int_x^\infty \phi(s)\mathrm{d}s = \lim_{\delta\to 0}\Big[\phi(x)\delta + \phi(x+\delta)\delta + \phi(x+2\delta)\delta + \dots\Big].$$

Using the expression for $\phi(\cdot)$ we get

$$\phi(x+\delta) = \phi(x)\mathrm{e}^{-x\delta}\mathrm{e}^{-\delta^2/2}$$

and hence we may write the denominator in (3.2.2) as

$$x\phi(x)\lim_{\delta\to 0}\Big[1 + \mathrm{e}^{-x\delta}\mathrm{e}^{-\delta^2/2} + \mathrm{e}^{-2x\delta}\mathrm{e}^{-4\delta^2/2} + \mathrm{e}^{-3x\delta}\mathrm{e}^{-9\delta^2/2} + \dots\Big]\delta.$$

This expressions $\mathrm{e}^{-x\delta}$, $\mathrm{e}^{-2x\delta}$ and so on for a geometric sequence with sum $1/(1-\mathrm{e}^{-x\delta}) \approx 1/x\delta$ as $x\delta \to 0$. Accordingly, we let $\delta \to 0$ and $x \to \infty$ in such as way that $x\delta \to 0$. We have to show that the other terms $\mathrm{e}^{n^2\delta^2/2}$ do not "disturb" the geometric sequence too much.

Now find n^*, starting of which the residual sum on the geometric sequence $1 + \mathrm{e}^{-x\delta} + \mathrm{e}^{-2x\delta} + \dots$ is less than $\varepsilon > 0$:

$$\sum_{n=n^*}^{\infty} \mathrm{e}^{-nx\delta} < \varepsilon \qquad \Rightarrow \qquad n^* > -\frac{\log\varepsilon + \log(1-\mathrm{e}^{-x\delta})}{x\delta}$$

choose $n^{**} > -\frac{\log\varepsilon}{x\delta} + 1 > n^*$. Now find

$$\exp\frac{n^{**2}\delta^2}{2} = \exp\left(\frac{\log^2\varepsilon}{2x^2} + \delta\frac{\log\varepsilon}{x} + \frac{\delta^2}{2}\right).$$

Because all the terms in parenthesis converge to as $x \to \infty$ and $\delta \to 0$, the exponent converges to 1. Hence, at the limit, the Rieman sum is solely determined by the geometric sequence, and we have denominator in (3.2.2) equal to $\phi(x)$. □

3.3 Differentiation

3.3.1 Simple Derivatives

$$\frac{\partial}{\partial x} a^x = a^x \log a \tag{3.3.1}$$

$$\tan \phi' = \frac{1}{\cos^2 \phi} \tag{3.3.2}$$

$$\arctan x' = \frac{1}{1 + x^2} \tag{3.3.3}$$

3.3.2 Directional Derivative

$$f'(\boldsymbol{x}; \boldsymbol{u}) = \lim_{h \to 0} \frac{f(\boldsymbol{x} + h\boldsymbol{u}) - f(\boldsymbol{x})}{h} \tag{3.3.4}$$

3.3.3 Jacobian Matrix

$$\mathcal{J} = \begin{bmatrix} \dfrac{\partial \mathbf{f}}{\partial x_1} & \cdots & \dfrac{\partial \mathbf{f}}{\partial x_n} \end{bmatrix} = \frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)} = \boldsymbol{D} f(\boldsymbol{x}) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix} \tag{3.3.5}$$

3.3.4 Ümbrikuteoreem (*envelope theorem*)

Olgu M defineeritud kui optimum funktsioonist f:

$$M(a) = \max_x f(x, a), \tag{3.3.6}$$

kus a on parameeter. Siis

$$\frac{\mathrm{d}M(a)}{\mathrm{d}a} = \left. \frac{\partial f(x^*, a)}{\partial a} \right|_{x^* = x(a)}. \tag{3.3.7}$$

3.3.5 Normal Density Related Derivatives

One-Dimensional Case

$$\frac{\mathrm{d}}{\mathrm{d}x} \Phi(-x) = \phi(x) \tag{3.3.8}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \frac{1}{\sigma} \phi\left(\frac{x - \mu}{\sigma}\right) = -\frac{1}{\sigma} \left(\frac{x - \mu}{\sigma}\right) \phi\left(\frac{x - \mu}{\sigma}\right) \tag{3.3.9}$$

$$\frac{\mathrm{d}}{\mathrm{d}\sigma} \frac{1}{\sigma} \phi\left(\frac{x - \mu}{\sigma}\right) = \frac{1}{\sigma^2} \left[\left(\frac{x - \mu}{\sigma}\right)^2 - 1\right] \phi\left(\frac{x - \mu}{\sigma}\right) \tag{3.3.10}$$

Multi-Dimensional Case Let $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$ be the 2-dimensional variance-covariance matrix and $\phi(\cdot,\cdot)$ the 2-D normal density, defined in (5.5.2):

$$\frac{\partial}{\partial x_1}\phi(\boldsymbol{x},\boldsymbol{\Sigma}) = \phi(\boldsymbol{x},\boldsymbol{\Sigma})\frac{\rho x_2 - x_1}{1-\rho^2} \tag{3.3.11}$$

$$\begin{aligned}\frac{\partial}{\partial \rho}\phi(\boldsymbol{x},\boldsymbol{\Sigma}) &= \phi(\boldsymbol{x},\boldsymbol{\Sigma})\left[\frac{\rho}{1-\rho^2}\left(1 - \boldsymbol{x}'\boldsymbol{\Sigma}^{-1}\boldsymbol{x}\right) + \frac{x_1x_2}{1-\rho^2}\right] = \\ &= \phi(\boldsymbol{x},\boldsymbol{\Sigma})\left[\frac{\rho}{1-\rho^2} - \frac{\rho}{(1-\rho^2)^2}(x_1^2 - 2\rho x_1x_2 + x_2^2) + \frac{x_1x_2}{1-\rho^2}\right]\end{aligned} \tag{3.3.12}$$

$$\frac{\partial^2}{\partial x_1 \partial x_2}\phi(\boldsymbol{x},\boldsymbol{\Sigma}) = \phi(\boldsymbol{x},\boldsymbol{\Sigma})\left[\frac{(x_1-\rho x_2)(x_2 - \rho x_1)}{(1-\rho^2)^2} + \frac{\rho}{1-\rho^2}\right] \tag{3.3.13}$$

3.3.6 Derivatives of Gamma Function

$$\Gamma'(x) = \psi(x)\Gamma(x) \tag{3.3.14}$$

$$\Gamma''(x) = \Gamma(x)\left[\psi'(x) + \psi^2(x)\right], \tag{3.3.15}$$

where $\psi(x)$ is the digamma function.

3.3.7 Differentiation of Sums

$$\frac{\partial}{\partial x_i}\sum_j x_j = 1 \tag{3.3.16}$$

$$\frac{\partial}{\partial x_i}\left(\sum_j x_j\right)^2 = 2\sum_j x_j \tag{3.3.17}$$

3.3.8 General Differentiation Rules

Pöördfunktsiooni tuletis

$$\frac{\mathrm{d}}{\mathrm{d}x}f^{-1}(x) = \left.\frac{1}{\frac{\mathrm{d}}{\mathrm{d}x}f(x)}\right|_{x=f^{-1}(y)} = \frac{1}{f'[f^{-1}(y)]} \tag{3.3.18}$$

Mitmekihilise (liit-) funktsiooni tuletis Olgu $v = v(b)$ ja $f(v) = f(v(b)) = g(b)$. Tuletised:

$$\frac{\partial g}{\partial b} = \frac{\partial f}{\partial v}\cdot\frac{\partial v}{\partial b} \equiv f'(v(b))\cdot v'(b) \tag{3.3.19}$$

$$\frac{\partial^2 g}{\partial b^2} = \frac{\partial^2 g}{\partial v^2}\left(\frac{\partial v}{\partial b}\right)^2 + \frac{\partial f}{\partial v}\frac{\partial^2 v}{\partial b^2} \tag{3.3.20}$$

3.4 Integration

3.4.1 Simple Integrals

$$\int x^{\alpha} = \frac{1}{\alpha+1} x^{\alpha+1}, \quad \alpha \neq -1 \tag{3.4.1}$$

Leibnitz' Rule

$$\frac{\partial}{\partial x} \int_{a(x)}^{b(x)} f(y)\, \mathrm{d}y = f[b(x)]b'(x) - f[a(x)]a'(x) \tag{3.4.2}$$

$$\begin{aligned} \frac{\partial}{\partial x} \int_{a(x)}^{b(x)} f(y,x)\, \mathrm{d}y &= f[b(x),x]b'(x) - f[a(x),x]a'(x) + \\ &+ \int_{a(x)}^{b(x)} \frac{\partial}{\partial x} f(y,x)\, \mathrm{d}y \end{aligned} \tag{3.4.3}$$

3.4.2 Integrals related to probability distributions

Normal distribution Let

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2}. \tag{3.4.4}$$

To prove the following, simply integrate $\phi(\cdot)$, in most cases by parts.

$$\int \phi(x)\, \mathrm{d}x \equiv \Phi(x) \tag{3.4.5}$$

$$\int_a^b \phi\left(\frac{x-\mu}{\sigma}\right) \mathrm{d}x = \sigma\Phi\left(\frac{b-\mu}{\sigma}\right) - \sigma\Phi\left(\frac{a-\mu}{\sigma}\right) \tag{3.4.6}$$

$$\int \phi^2(x)\, \mathrm{d}x = \frac{1}{2\pi}\Phi(\sqrt{2}x) \tag{3.4.7}$$

$$\int x\phi(x)\, \mathrm{d}x = -\phi(x) \tag{3.4.8}$$

$$\int_{-\infty}^{\infty} x\phi\left(\frac{x-\mu}{\sigma}\right) \mathrm{d}x = \sigma\mu \tag{3.4.9}$$

$$\int x\phi\left(\frac{x-\mu}{\sigma}\right) \mathrm{d}x = \sigma\mu\Phi\left(\frac{x-\mu}{\sigma}\right) - \sigma^2\phi\left(\frac{x-\mu}{\sigma}\right) + C \tag{3.4.10}$$

$$\begin{aligned} \int_a^b x\phi\left(\frac{x-\mu}{\sigma}\right) \mathrm{d}x &= \sigma\mu\left[\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\right] + \\ &+ \sigma^2\left[\phi\left(\frac{a-\mu}{\sigma}\right) - \phi\left(\frac{b-\mu}{\sigma}\right)\right] \end{aligned} \tag{3.4.11}$$

$$\int x^2\phi(x)\, \mathrm{d}x = -x\phi(x) + \Phi(x) \tag{3.4.12}$$

$$\int_a^b x^2\phi(x)\, \mathrm{d}x = a\phi(a) - b\phi(b) + \Phi(b) - \Phi(a) \tag{3.4.13}$$

$$\int x^2\phi\left(\frac{x}{\sigma}\right) \mathrm{d}x = -\sigma^2 x\phi\left(\frac{x}{\sigma}\right) + \sigma^3\Phi\left(\frac{x}{\sigma}\right) \tag{3.4.14}$$

$$\int_{-\infty}^{\infty} x^2\phi\left(\frac{x-\mu}{\sigma}\right)\mathrm{d}x = \sigma(\mu^2+\sigma^2) \tag{3.4.15}$$

$$\int_{-\infty}^{\infty} \mathrm{e}^{\alpha x}\phi\left(\frac{x-\mu}{\sigma}\right)\mathrm{d}x = \mathrm{e}^{\alpha\mu+\frac{1}{2}\sigma^2\alpha^2} \tag{3.4.16}$$

$$\int_{s}^{t} \phi(u)\log\phi(u)\,\mathrm{d}u = \frac{1}{2}\left[\Phi(s)-\Phi(t)\right](1+\log 2\pi) + \frac{1}{2}\left[t\phi(t)-s\phi(s)\right] \tag{3.4.17}$$

$$\begin{aligned}\int_{s}^{t} \phi(u)\log\phi(u-\mu)\,\mathrm{d}u &= \frac{1}{2}\left[\Phi(s)-\Phi(t)\right]\left(1+\log 2\pi+\mu^2\right)+\\ &\quad+\frac{1}{2}\left[t\phi(t)-s\phi(s)\right]+\mu\left[\phi(s)-\phi(t)\right]\end{aligned} \tag{3.4.18}$$

$$\int x\phi^2(x)\,\mathrm{d}x = -\frac{1}{2}\phi^2(x) \tag{3.4.19}$$

$$\int \phi(x)\Phi(x)\,\mathrm{d}x = \frac{1}{2}\Phi^2(x) \tag{3.4.20}$$

$$\int x\phi(x)\Phi(x)\,\mathrm{d}x = -\phi(x)\Phi(x) + \frac{1}{2\sqrt{\pi}}\Phi(\sqrt{2}\,x) \tag{3.4.21}$$

$$\int x^2\phi(x)\Phi(x)\,\mathrm{d}x = \frac{1}{2}\Phi^2(x) - x\phi(x)\Phi(x) - \frac{1}{2}\phi^2(x) \tag{3.4.22}$$

Log-normal density Let $f(\cdot)$ be the log-normal density.

$$\begin{aligned}\int_{a}^{\infty} xf(x)\,\mathrm{d}x &= \int_{a}^{b} \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{\log x-\mu}{\sigma}\right)^2\right]\mathrm{d}x =\\ &= \frac{1-\Phi\left(\frac{\log a-\mu-\sigma^2}{\sigma}\right)}{1-\Phi\left(\frac{\log a-\mu}{\sigma}\right)}\mathrm{e}^{\mu+\frac{1}{2}\sigma^2}\end{aligned} \tag{3.4.23}$$

$$\begin{aligned}\int_{a}^{b} xf(x)\,\mathrm{d}x &= \int_{a}^{b} \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{\log x-\mu}{\sigma}\right)^2\right]\mathrm{d}x =\\ &= \mathrm{e}^{\mu+\frac{1}{2}\sigma^2}\left[-\Phi\left(\frac{\log a-\mu-\sigma^2}{\sigma}\right)+\Phi\left(\frac{\log b-\mu-\sigma^2}{\sigma}\right)\right]\end{aligned} \tag{3.4.24}$$

3.4.3 Other integrals

$$\int_{0}^{t} \mathrm{e}^{-rT}\,\mathrm{d}T = \frac{1}{r}\left[1-\mathrm{e}^{-rt}\right] \tag{3.4.25}$$

$$\int \frac{\mathrm{d}x}{(p+q^{ax})^2}\,\mathrm{d}x = \frac{x}{p^2} + \frac{1}{ap(p+q\mathrm{e}^{ax})} - \frac{1}{ap^2}\ln(p+q\mathrm{e}^{ax}) \tag{3.4.26}$$

$$\int \log x\,\mathrm{d}x = x\log x - x \tag{3.4.27}$$

$$\int \mathrm{e}^{-ax}\mathrm{d}x = \frac{1}{a}\left[1-\mathrm{e}^{-ax}\right] \tag{3.4.28}$$

$$\int x\mathrm{e}^x \mathrm{d}x = x\mathrm{e}^x - \mathrm{e}^x \tag{3.4.29}$$

$$\int x\mathrm{e}^{-x} \mathrm{d}x = -x\mathrm{e}^{-x} - \mathrm{e}^{-x} \tag{3.4.30}$$

$$\int x\mathrm{e}^{\alpha x}\,\mathrm{d}x = \frac{x}{\alpha}\mathrm{e}^{\alpha x} - \frac{1}{\alpha^2}\mathrm{e}^{\alpha x} \tag{3.4.31}$$

$$\int_a^b x\mathrm{e}^{\alpha x}\,\mathrm{d}x = \frac{1}{\alpha}\left[\mathrm{e}^{\alpha b}\left(b - \frac{1}{\alpha}\right) - \mathrm{e}^{\alpha a}\left(a - \frac{1}{\alpha}\right)\right] \tag{3.4.32}$$

$$\int x^2\mathrm{e}^{\alpha a}\,\mathrm{d}x = \frac{1}{\alpha}x^2\mathrm{e}^{\alpha x} - \frac{2}{\alpha}\int x\mathrm{e}^{\alpha x}\,\mathrm{d}x \tag{3.4.33}$$

$$\int_a^b x^2\mathrm{e}^{\alpha a}\,\mathrm{d}x = \frac{1}{\alpha}\left[\mathrm{e}^{\alpha b}\left(b^2 - \frac{2b}{\alpha} + \frac{2}{\alpha^2}\right) - \mathrm{e}^{\alpha a}\left(a^2 - \frac{2a}{\alpha} + \frac{2}{\alpha^2}\right)\right] \tag{3.4.34}$$

$$\int_0^\infty x^\alpha \mathrm{e}^{-\beta x}\mathrm{d}x = \frac{\Gamma(\alpha+1)}{\beta^{\alpha+1}} \tag{3.4.35}$$

Let $f(\cdot)$ be a distribution function and $\bar{F}(\cdot)$ the corresponding survival function:

$$\int_c^b \left[f(x)\int_c^x w(y)\,\mathrm{d}y\right]\mathrm{d}x = \int_c^b \bar{F}(x)w(x)\,\mathrm{d}x \tag{3.4.36}$$

Euleri konstant

$$\int_0^\infty e^{-z}\log z\,\mathrm{d}z = c \approx -0,5772 \tag{3.4.37}$$

3.4.4 Üldised integreerimise reeglid

Muutuja vahetus integraali all Olgu vaja üle minna muutujatelt $(x_1 \ldots, x_N)$ muutujatele $(y_1, \ldots, y_N)$. Sel juhul

$$\begin{aligned}\int_V f(x_1,\ldots,x_N)\mathrm{d}x_1\ldots\mathrm{d}x_N &= \\ = \int_V f[y_1(x_1,\ldots,x_N),\ldots,y_N(x_1,\ldots,x_N)]\frac{\mathrm{d}y_1\ldots\mathrm{d}y_N}{|\mathcal{J}|} &= \\ = \int_V g(y_1,\ldots,y_N)\frac{\mathrm{d}y_1\ldots\mathrm{d}y_N}{|\mathcal{J}|}&,\end{aligned} \tag{3.4.38}$$

kus

$$|\mathcal{J}| = \left|\frac{\partial(y_1,\ldots,y_N)}{\partial(x_1,\ldots,x_N)}\right| \tag{3.4.39}$$

on koordinaatteisenduse jakobjaani absoluutväärtus.

3.4.5 Analüütilised funktsioonid

Gammafunktsioon

$$\Gamma(p) = \int_0^\infty \lambda^{p-1}e^{-\lambda}\mathrm{d}\lambda \tag{3.4.40}$$

Gammafunktsiooni omadus:

$$\Gamma(p+1) = p\Gamma(p) = p! \tag{3.4.41}$$

Tøestus: integreeri ositi.

Digamma funktsioon

$$\psi(\alpha) = \frac{d \log \Gamma(\alpha)}{d\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^{-x} \log x dx. \tag{3.4.42}$$

Digamma omadus:

$$\psi(\alpha+1) = \frac{1}{\alpha} + \psi(\alpha). \tag{3.4.43}$$

Digamma arvväärtused:

$$\psi(1) = -0,5772 \qquad \psi(2) = 0,4228 \tag{3.4.44}$$

Beetafunktsioon

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} = \int_0^1 x^{p-1}(1-x)^{q-1} dx, \tag{3.4.45}$$

kusjuures $p > 0$ ja $q > 0$.

3.4.6 Laplace'i teisendus

Laplace'i teisendus juhusliku muutuja X jaotusfunktsioonist on

$$L_f(s) = \mathbb{E}\, \mathrm{e}^{-sX} = \int \mathrm{e}^{-sx} dF_X(x). \tag{3.4.46}$$

Laplace'i teisendus on sama mis momendifunktsioon.

3.4.7 Numbriline integreerimine

Monte-Carlo integraal Olgu vaja leida

$$I = \int_a^b f(x)\,\mathrm{d}x = (b-a) \int_a^b f(x) \frac{1}{b-a}\,\mathrm{d}x = (b-a)\,\mathbb{E}[f(X)], \tag{3.4.47}$$

kus $X \sim U(a,b)$. Valimis suurusega N olgu $x_1, \dots x_N \sim i.i.d U(a,b)$. Siis integraali hinnanguks on funktsiooni väärtuste keskmine ja veahinnanguks tema standardhälve valimis:

$$\hat{I}_N = (b-a) \frac{1}{N} \sum_{i=1}^N f(x_i) \tag{3.4.48}$$

$$\widehat{\operatorname{Var} \hat{I}_N} = \frac{(b-a)^2}{N} \frac{1}{N} \sum_{i=1}^N \left[f(x_i) - \frac{1}{N} \hat{I}_N \right]^2 \tag{3.4.49}$$

3.5 Differential Equations

Let c be a constant.

$$\begin{aligned} y(x)' + Py &= Q \\ y(x) &= \frac{Q}{P} + c\mathrm{e}^{-Px} \end{aligned} \tag{3.5.1}$$

$$\begin{aligned} y(x)' + P(x)y &= Q(x) \\ y(x) &= \mathrm{e}^{-\int P(x)\,\mathrm{d}x} \int Q(x)\mathrm{e}^{\int P(x)\,\mathrm{d}x}\,\mathrm{d}x + c\mathrm{e}^{-\int P(x)\,\mathrm{d}x} \end{aligned} \tag{3.5.2}$$

3.6 Optimization

3.6.1 Second-order maximum/minimum conditions for constrained optimisation

The problem is

$$\begin{aligned} \max z &= f(x_1, x_2, \dots, x_n) \\ \text{s.t. } g(x_1, x_2, \dots, x_n) &= 0. \end{aligned} \tag{3.6.1}$$

Corresponding Lagrangian is

$$Z = f(x_1, x_2, \dots, x_n) - \lambda g(x_1, x_2, \dots, x_n). \tag{3.6.2}$$

Corresponding *bordered Hessian* is:

$$\left|\bar{\mathsf{H}}\right| = \begin{vmatrix} 0 & g_1 & g_2 & \dots & g_n \\ g_1 & Z_{11} & Z_{12} & \dots & Z_{1n} \\ g_2 & Z_{21} & Z_{22} & \dots & Z_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_n & Z_{n1} & Z_{n2} & \dots & Z_{nn} \end{vmatrix} \tag{3.6.3}$$

and successive *principal minors* are:

$$\left|\bar{\mathsf{H}}_2\right| = \begin{vmatrix} 0 & g_1 & g_2 \\ g_1 & Z_{11} & Z_{12} \\ g_2 & Z_{21} & Z_{22} \end{vmatrix} \qquad \left|\bar{\mathsf{H}}_3\right| = \begin{vmatrix} 0 & g_1 & g_2 & g_3 \\ g_1 & Z_{11} & Z_{12} & Z_{13} \\ g_2 & Z_{21} & Z_{22} & Z_{23} \\ g_3 & Z_{31} & Z_{32} & Z_{33} \end{vmatrix} \qquad \dots \qquad \left|\bar{\mathsf{H}}_n\right| = \left|\bar{\mathsf{H}}\right| \tag{3.6.4}$$

The second derivative $\mathrm{d}^2 z$ is positive definite iff

$$\bar{\mathsf{H}}_2 < 0, \quad \bar{\mathsf{H}}_3 < 0, \quad \dots, \quad \bar{\mathsf{H}}_n < 0,$$

and negavitve definite iff

$$\bar{\mathsf{H}}_2 > 0, \quad \bar{\mathsf{H}}_3 < 0, \quad \bar{\mathsf{H}}_n > 0, \quad \dots$$

Note that $\bar{\mathsf{H}}_1$ is always negative.
Proof: Chiang (1984).

3.6.2 Optimal control (dynamic optimisation)

The problem is:

$$\begin{aligned} &\max_u \int_0^T F(t,y,u)\,\mathrm{d}t \\ &\text{s.t. } \dot{y} = f(t,y,u) \\ &y(0) = A \qquad y(T) \text{ free.} \end{aligned} \tag{3.6.5}$$

Corresponding *Hamiltonian* is

$$H(t,y,u,\lambda) \equiv F(t,y,u) + \lambda(t) f(t,y,u). \tag{3.6.6}$$

The first order conditions for optimum are

1. $\max_u H(t,y,u,\lambda) \quad \forall t \in [0,T]$ or, less generally, $\frac{\partial H}{\partial t} = 0$ (optimality condition).
2. $\dot{y} = \frac{\partial H}{\partial \lambda}$ (equation of motion for y).
3. $\dot{\lambda} = -\frac{\partial H}{\partial y}$ (equation of motion for λ).
4. $\lambda(T) = 0$ (transversality condition).

Proof: Miller (1979)

3.6.3 Newton-Raphson maximization

Non-linear continuous function of N-dimensional parameter can, under suitable assumptions, be approximated as N-dimensional parabola. When running non-linear maximization, we may approximate the function in this way at the initial value of the parameter vector. The maximum of the approximation can be used as the initial value for the next step.

Let us maximise a function $l(\boldsymbol{\vartheta})$ where $\boldsymbol{\vartheta}$ is a N-dimesional parameter vector. Let the initial value of the parameter be $\boldsymbol{\vartheta}_0$. From Taylor's approximation:

$$l(\boldsymbol{\vartheta}) \approx l(\boldsymbol{\vartheta}_0) + \left.\frac{\partial l(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}}\right|_{\boldsymbol{\vartheta}=\boldsymbol{\vartheta}_0} (\boldsymbol{\vartheta} - \boldsymbol{\vartheta}_0) + \frac{1}{2}(\boldsymbol{\vartheta} - \boldsymbol{\vartheta}_0)' \left.\frac{\partial^2 l(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta} \partial \boldsymbol{\vartheta}'}\right|_{\boldsymbol{\vartheta}=\boldsymbol{\vartheta}_0} (\boldsymbol{\vartheta} - \boldsymbol{\vartheta}_0) \tag{3.6.7}$$

At the maximum $\partial l(\boldsymbol{\vartheta})/\partial \boldsymbol{\vartheta} = 0$ and hence the parameter value at the maximum (the initial value for the next iteration):

$$\boldsymbol{\vartheta}_1 = \boldsymbol{\vartheta}_0 - \left[\left.\frac{\partial^2 l(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta} \partial \boldsymbol{\vartheta}'}\right|_{\boldsymbol{\vartheta}=\boldsymbol{\vartheta}_0} \right]^{-1} \left.\frac{\partial l(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}}\right|_{\boldsymbol{\vartheta}=\boldsymbol{\vartheta}_0} \tag{3.6.8}$$

The algorith requires either programming the analytical Hessian matrix $\frac{\partial^2 l(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta} \partial \boldsymbol{\vartheta}'}$, or calculating the Hessian matrix by numeric differentiation. The first way may be complicated, the latter one slow and subject to numerical errors.

BHHH maximization BHHH is a particular version of Newton-Raphson algorithm, suitable for maximizing log-likelihood function only. BHHH uses the information equlity for approximating the Hessian:

$$\mathbb{E}\left[\frac{\partial^2 l(\vartheta)}{\partial\vartheta\partial\vartheta'}\right]_{\vartheta=\vartheta_0} = -\,\mathbb{E}\left[\left.\frac{\partial l(\vartheta)}{\partial\vartheta'}\right|_{\vartheta=\vartheta_0}\left.\frac{\partial l(\vartheta)}{\partial\vartheta}\right|_{\vartheta=\vartheta_0}\right] \tag{3.6.9}$$

This algorithm does not require Hessian matrix (this is approximated). However, it typically requires around 10 times more iterations for convergence as the approximation may be quite imprecise when initial values are far off the target. Note also that while the estimates are exactly the same as in the case of NR algorithm, the standard errors may be different on a finite sample (Calzolari and Fiorentini, 1993).

4 Algebra

4.1 Mõisted

proper subset A on B *proper subset* kui $A \subseteq B$ kui $B \nsubseteq A$.

proper subspace Kui A ja B ja A on B *proper subset*. Näiteks tasandi tõeline (lineaarne) alamruum on sirge.

4.2 Tehted hulkadega

$$\left(\bigcap_{A \in F} A\right)^C = \bigcup_{A \in F} A^C \tag{4.2.1}$$

$$\left(\bigcup_{A \in F} A\right)^C = \bigcap_{A \in F} A^C \tag{4.2.2}$$

$$A \setminus B = A \cap B^C \tag{4.2.3}$$

4.3 Simple Algebra

Binomial Theorem

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i \tag{4.3.1}$$

where $\binom{n}{i}$ is the *binomial coefficient*, number of distinct combinations of i elements out of n elements in total: $\binom{n}{i} \equiv C_n^i = \frac{n!}{(n-i)!i!}$.

$$\sum_{i=0}^{n} \binom{n}{i} \mathrm{e}^{\alpha i} = (1+\mathrm{e}^{\alpha})^n \tag{4.3.2}$$

$$\sum_{i=0}^{n} i \binom{n}{i} \mathrm{e}^{\alpha i} = n\mathrm{e}^{\alpha}(1+\mathrm{e}^{\alpha})^{n-1} \tag{4.3.3}$$

Geomeetrilise jada summa

$$S = 1 + q + q^2 + q^3 + \ldots = \frac{1}{1-q}. \tag{4.3.4}$$

Tõestus: kirjuta välja qS, lahuta ja avalda S. Märkus: kui jada on kujul $S' = q + q^2 + q^3 + \ldots$, siis $S' = qS$. Oluline erijuht kui $q = \frac{1}{1+r}$:

$$S = 1 + \frac{1}{1+r} + \frac{1}{(1+r)^2} + \ldots = 1 + \frac{1}{r}. \tag{4.3.5}$$

$$\sum_{i=0}^{\infty} i p^i = \frac{p}{(1-p)^2} \tag{4.3.6}$$

Tõestus: kui S on antud summa, siis avalda $S - pS \ldots$

Eksponent piirväärtusena

$$\lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n = e^a. \tag{4.3.7}$$

Tõestus: arenda Newtoni binoomvalemiga ritta, arvesta (4.3.8).

faktoriaalide jagatis

$$\lim_{n\to\infty}\frac{n!}{(n-m)!} = n^m \tag{4.3.8}$$

Märkus: siin on eeldatud, et $m \not\to \infty$. Jaga läbi, arvesta, et $n-1 \approx n$. Seose erijuht:

$$\lim_{n\to\infty}\frac{n}{m} = \lim_{n\to\infty}\frac{n!}{(n-m)!m!} = \frac{n^m}{m!}. \tag{4.3.9}$$

4.4 Taylori rida

Taylori rida Iga funktsiooni võib punkti x_0 ümbruses esitada astmereana:

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2}f''(x_0)(x-x_0)^2 + \ldots = \\ &= \sum_{i=0}^{\infty} f^{(i)}(x_0)\frac{(x-x_0)^i}{i!}. \end{aligned} \tag{4.4.1}$$

Tõestus: kirjuta samasugune astmerida tundmatute kordajatega välja, võrruta $f(x-x_0)$-ga ja võta järjest tuletisi. Taylori rea erijuht, kui $x_0 = 0$ on McLaureni rida:

$$f(x) = f(0) + f'(0)x + \frac{1}{2}f''(0)x^2 + \ldots = \sum_{i=0}^{\infty} f^{(i)}(0)\frac{x^i}{i!}. \tag{4.4.2}$$

Eksponendi astmerida

$$e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \ldots = \sum_{i=0}^{\infty}\frac{x^i}{i!}. \tag{4.4.3}$$

Tõestus: arenda e^x Taylori ritta.

Eksponendi piirväärtus

$$\lim_{x\to 0} e^x = 1 + x. \tag{4.4.4}$$

Tõestus: eksponendi astmereast. Märkus: piirväärtus $1+x$ on kõige tavalisem, mida on vaja kasutada. Olenevalt ülesandest tuleb arvestada rohkem (või ka vähem) astmerea liikmeid.

Logaritmi astmerida

$$\ln x = (x-1) - \frac{(x-1)^2}{2!} + \frac{(x-1)^3}{3!} - \frac{(x-1)^4}{4!} + \ldots \tag{4.4.5}$$

Tõestus: arenda Taylori ritta $x_0 = 1$ ümbruses.

4.5 Summade astmed

$$\left(\sum_{i=1}^{N} x_i\right)^2 = \sum_{i=1}^{N} x_i^2 + \sum_{\substack{i;j=1\\ i\neq j}}^{N} x_i x_j \tag{4.5.1}$$

$$\left(\sum_{i=1}^{N} x_i\right)^3 = \sum_{i=1}^{N} x_i^3 + 3\sum_{\substack{i;j=1\\ i\neq j}}^{N} x_i x_j^2 + \sum_{\substack{i;j;k=1\\ i\neq j;j\neq k;k\neq i}}^{N} x_i x_j x_k \tag{4.5.2}$$

$$\begin{aligned}\left(\sum_{i=1}^{N} x_i\right)^4 &= \sum_{i=1}^{N} x_i^4 + 4\sum_{\substack{i;j=1\\ i\neq j}}^{N} x_i x_j^3 + 3\sum_{\substack{i;j=1\\ i\neq j}}^{N} x_i^2 x_j^2 + 6\sum_{\substack{i;j;k=1\\ i\neq j;j\neq k;k\neq i}}^{N} x_i x_j x_k^2 + \\ &+ \sum_{\substack{i;j;k;l=1\\ i;j;k;l\neq}}^{N} x_i x_j x_k x_l\end{aligned} \tag{4.5.3}$$

Ühekordsetes summades on N liiget, kahekordsetes $N(N-1)$, kolmekordsetes $N(N-1)(N-2)$ ning neljakordses $N(N-1)(N-2)(N-3)$.

Tuletuskäik lähtub viimasel juhul niisugustest mɵtetest:

- Kui komponentide indeksid ei tohi olla vɵrdsed, siis on järgmist komponenti vɵimalik valida ühe vɵrra vähem
- Esimesel liikmel vɵetakse sisse kɵik komponendid, seega on C_0^4 varianti.
- Teisel liikmel on kaks komponenti (x_i ja x_j), ühte vɵetakse kolm korda, teist korra. Seega tuleb valida üks, mis iga kord välja jäetakse. Seega C_1^4 vɵimalust.
- Kolmandal liikmel valitakse mɵlemad komponendid kahe kaupa. Kokku on $C_2^4 = 6$ vɵimalust kahe kaupa valida, kuna aga pole vahet kumb komponentidest on kumb, siis jääb järele pool nendest.
- Neljandal liikmel on kolm komponenti, ruutliikme valimiseks on $C_2^4 = 6$ vɵimalust. Kuna teised liikmed on esimeses astmes, siis on küll vahe, kumbad me välja valime. Jääb 6.
- Viimane, kɵiki üks kord, $C_4^4 = 1$.

Kui $X \sim i.i.d$, siis

$$\mathbb{E}\left(\sum^{N} x_i\right)^2 = N\,\mathbb{E}\,X^2 + N(N-1)(\mathbb{E}\,X)^2 = N^2(\mathbb{E}\,X)^2 + N\mathrm{Var}\,X \tag{4.5.4}$$

4.6 Maatriksalgebra

$$\mathsf{A}^{-1} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{|A|}\begin{bmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{bmatrix} \tag{4.6.1}$$

$$\begin{bmatrix} \mathsf{A} & \mathsf{B} \\ \mathsf{C} & \mathsf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathsf{E}^{-1} & -\mathsf{E}^{-1}\mathsf{B}\mathsf{D}^{-1} \\ -\mathsf{D}^{-1}\mathsf{C}\mathsf{E}^{-1} & \mathsf{F}^{-1} \end{bmatrix}, \quad \text{kus} \tag{4.6.2}$$

$$\begin{aligned} \mathsf{E} &= \mathsf{A} - \mathsf{B}\mathsf{D}^{-1}\mathsf{C} \\ \mathsf{E}^{-1} &= \mathsf{A}^{-1} + \mathsf{A}^{-1}\mathsf{B}\mathsf{F}^{-1}\mathsf{C}\mathsf{A}^{-1} \\ \mathsf{F} &= \mathsf{D} - \mathsf{C}\mathsf{A}^{-1}\mathsf{B} \\ \mathsf{F}^{-1} &= \mathsf{D}^{-1}\mathsf{C}\mathsf{E}^{-1}\mathsf{B}\mathsf{D}^{-1} \end{aligned} \tag{4.6.3}$$

4.7 Differentiation of Matrices and Vectors

Definition: let λ be a scalar, $\boldsymbol{x}$ and $\boldsymbol{\beta}$ $K \times 1$ vectors, and A a $K \times K$ matrix. Define:

$$\frac{\partial \lambda}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial}{\partial x_1}\lambda \\ \dots \\ \frac{\partial}{\partial x_K}\lambda \end{bmatrix} \tag{4.7.1}$$

More results:

$$\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{x}'} = \begin{bmatrix} \frac{\partial x_1}{\partial x_1} & \frac{\partial x_1}{\partial x_2} & \dots & \frac{\partial x_1}{\partial x_K} \\ \frac{\partial x_2}{\partial x_1} & \frac{\partial x_2}{\partial x_2} & \dots & \frac{\partial x_2}{\partial x_K} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_K}{\partial x_1} & \frac{\partial x_K}{\partial x_2} & \dots & \frac{\partial x_K}{\partial x_K} \end{bmatrix} = \mathsf{I} \qquad \frac{\partial \boldsymbol{x}'}{\partial \boldsymbol{x}} = \mathsf{I} \tag{4.7.2}$$

$$\frac{\partial \boldsymbol{x}'\boldsymbol{\beta}}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial}{\partial x_1}(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_K x_K) \\ \frac{\partial}{\partial x_2}(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_K x_K) \\ \vdots \\ \frac{\partial}{\partial x_K}(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_K x_K) \end{bmatrix} = \boldsymbol{\beta} \tag{4.7.3}$$

$$\frac{\partial \mathsf{A}\boldsymbol{x}}{\partial \boldsymbol{x}'} = \mathsf{A} \qquad \frac{\partial \boldsymbol{x}'\mathsf{A}}{\partial \boldsymbol{x}} = \mathsf{A} \tag{4.7.4}$$

$$\frac{\partial \boldsymbol{x}'\boldsymbol{x}}{\partial \boldsymbol{x}} = \mathsf{I}\boldsymbol{x} + \boldsymbol{x}\mathsf{I} = 2\boldsymbol{x} \qquad \frac{\partial \boldsymbol{x}'\mathsf{A}\boldsymbol{x}}{\partial \boldsymbol{x}} = (\mathsf{A} + \mathsf{A}')\boldsymbol{x} \tag{4.7.5}$$

Let $\boldsymbol{f}(\boldsymbol{x})$ be a $M \times 1$ vector function and B a $M \times M$ matrix.

$$\frac{\partial \boldsymbol{f}(\boldsymbol{x})'\mathsf{B}\boldsymbol{f}(\boldsymbol{x})}{\partial \boldsymbol{x}} = [\boldsymbol{D}\, \boldsymbol{f}(\boldsymbol{x})]'\,(\mathsf{B} + \mathsf{B}')\boldsymbol{f}(\boldsymbol{x}), \tag{4.7.6}$$

where $\boldsymbol{D}\,\boldsymbol{f}(\boldsymbol{x})$ is Jacobian Matrix (3.3.3).

Oluline: ühte korrutamist ei tohi teiseks muuta. Näiteks kui avaldis sisaldab nii maatrikskorrutist kui skalaariga korrutamist (skalaariga korrutamine on põhimõtteliselt sama mis Kroneckeri korrutis $\otimes$), ei tohi endist skalaariga korrutamist diferentseerimise järel tõlgendada maatrikskorrutisena. Mis siis et skalaari asemel on nüüd maatriks:

$$\frac{\partial}{\partial \boldsymbol{\beta}'} [(\boldsymbol{\beta}' \boldsymbol{x}) \otimes \boldsymbol{y}] = \frac{\partial \boldsymbol{\beta}' \boldsymbol{x}}{\partial \boldsymbol{\beta}'} \otimes \boldsymbol{y} = \boldsymbol{x}' \otimes \boldsymbol{y} = \boldsymbol{y}\boldsymbol{x}'. \tag{4.7.7}$$

Skalaariga korrutamisel korrutatakse kõik maatriksi elemendid läbi sama skalaariga, seega pääle tuletise võtmist tuleb kõik vektori $\boldsymbol{y}$ elemendid läbi korrutada tuletisvektoriga $\boldsymbol{x}'$.

4.8 Vərratused

4.8.1 Hölder's inequality

Let X and Y be random variables.

$$\mathbb{E}\,|XY| \leq \left\{\mathbb{E}\left[|X|^{\frac{1}{\alpha}}\right]\right\}^{\alpha}\left\{\mathbb{E}\left[|X|^{\frac{1}{1-\alpha}}\right]\right\}^{1-\alpha} \tag{4.8.1}$$

4.8.2 Jensen Inequality

$$Ef(x) < f(Ex), \tag{4.8.2}$$

If $f(x)$ is concave.

4.8.3 Triangle Inequality

$$|x+y| \leq |x| + |y| \tag{4.8.3}$$

4.8.4 Cauchy-Schwartzi vərratus

$$| < x, y > | \leq ||x|| \cdot ||y|| \tag{4.8.4}$$

For sequences

$$\left(\sum a_i b_i\right)^2 \leq \left(\sum a_i\right)^2 \left(\sum b_i\right)^2. \tag{4.8.5}$$

For functions:

$$< x, y >= \int x \cdot y \, \mathrm{d}x \qquad ||x|| = \sqrt{< x, x >}$$

Vektorkujul:

$$(\boldsymbol{a} \cdot \boldsymbol{b})^2 \leq ||\boldsymbol{a}||^2 ||\boldsymbol{b}||^2 \tag{4.8.6}$$

ehk siis ka

$$\sum_i z_i z_i' \geq \frac{\sum_i a_i z_i \sum_i a_i z_i'}{\sum_i a_i^2} \tag{4.8.7}$$

4.8.5 Inequalities, containing exponent

Proof in most cases by analysing the corresponding function.

$$\mathrm{e}^a \geq a \tag{4.8.8}$$

$$1 - \mathrm{e}^{-a} < a \tag{4.8.9}$$

$$(1 - \mathrm{e}^{-a})\mathrm{e}^{-a} < a \tag{4.8.10}$$

$$(1 + a)\mathrm{e}^a \geq (1 + 2a) \tag{4.8.11}$$

$$(1 + a)\mathrm{e}^{-a} < 1 \qquad \text{if} \quad a > 0 \tag{4.8.12}$$

$$(a - 1)\mathrm{e}^a > -1 \qquad \text{if} \quad a > 0 \tag{4.8.13}$$

$$(1 + a^2)\mathrm{e}^{-a} < 1 \qquad \text{if} \quad a > 0 \tag{4.8.14}$$

$$\mathrm{e}^{-a} - \mathrm{e}^{-b} < -a + b \qquad \text{if} \quad 0 < a < b \tag{4.8.15}$$

$$\mathrm{e}^a - \mathrm{e}^b = (a - b) + \frac{1}{2}(a^2 - b^2) + \frac{1}{6}(a^3 - b^3) + \dots \tag{4.8.16}$$

$$\gtrless (a - b) \qquad \text{if} \quad a \gtrless b \tag{4.8.17}$$

$$ae^{-b} - be^{-a} \gtrless a - b \qquad \text{if} \quad a \gtrless b \tag{4.8.18}$$

$$a^2e^b - b^2e^a \gtrless a^2 - b^2 + ab(a-b) + \frac{1}{6}a^2b^2(b-a) + \frac{1}{24}a^2b^2(b^2-a^2) + \ldots$$
$$\text{if} \quad b \gtrless a \tag{4.8.19}$$

5 Probability and Statistics

5.1 Combinatorics

Combinations Number of combinations of selecting k items out of n where order does not matter:

$$\binom{n}{k} \equiv C_k^n = \frac{n!}{k!(n-k)!} \tag{5.1.1}$$

5.2 Distributions: General Concepts

5.2.1 Stohhastiline domineerimine

Olgu F ja G jaotusfunktsioonid. G esimest järku domineerib F-i kui $G(x) \leq F(x) \quad \forall x$ ehk $\bar{G}(x) \geq \bar{F}(x) \forall x$. Laias laastus annab G suuremad väärtused kui F.

5.2.2 Sɵltumatud juhuslikud muutujad

X ja Y on sɵltumatud $\Leftrightarrow f(x,y) = f_X(x) f_Y(y) \Leftrightarrow F(x,y) = F_X(x) F_Y(y)$.

5.2.3 Expectations

Let support of random variable X be $[a, b]$. Expectation of X

$$\mathbb{E}\, X = \int_a^b x \, \mathrm{d}F_X(x) \tag{5.2.1}$$

$$= a + \int_a^b \bar{F}_X(x) \, \mathrm{d}x. \tag{5.2.2}$$

Law of iterated expectations

$$\mathop{\mathbb{E}}_X [\mathbb{E}[Y|X]] = \mathbb{E}[Y] \tag{5.2.3}$$

5.2.4 Characteristic Function

Let X be a random variable. It's characteristic function:

$$\phi_X(t) = \mathbb{E}\, \mathrm{e}^{itX} \tag{5.2.4}$$

Properties: for independent random variables X_1, X_2

$$\phi_{X_1+X_2}(t) = \phi_{X_1}(t) \cdot \phi_{X_2}(t) \tag{5.2.5}$$

5.2.5 Momendifunktsioon (*moment generating function*)

MGF avaldub ühemɵɵtmelisel juhul:

$$M_x(s) = \mathbb{E}\, e^{sx} = \int e^{sx} f(x) dx. \tag{5.2.6}$$

Momendifunktsiooni omadused:

$$M_x'(0) = \mathbb{E}\,x \qquad M_x''(0) = \mathbb{E}\,x^2 \qquad M_x^{(n)}(0) = \mathbb{E}\,x^n. \tag{5.2.7}$$

Kasulik asi on ka $\log x$-i momendifunktsioon:

$$M_{\log x}(s) = \mathbb{E}\,\mathrm{e}^{s\log x} = \mathbb{E}\,x^s. \tag{5.2.8}$$

N-mõõtmelisel juhul avaldub MGF:

$$\begin{aligned}M(s_1, s_2, \ldots, s_N) &= \mathbb{E}\,e^{\sum^N s_i x_i} = \\ &= \int\cdots\int e^{s_1x_1}e^{s_2x_2}\ldots e^{s_Nx_N} f(x_1, x_2, \ldots, x_N)dx_1dx_2\ldots dx_N.\end{aligned} \tag{5.2.9}$$

5.2.6 Kumulandifunktsoon (*cumulant-generating function*)

KGF avaldub MGF-i kaudu:

$$K_x(s) = \log M(s) \qquad \text{või} \qquad K(s_1, s_2, \ldots, s_N) = \log M(s_1, s_2, \ldots, s_N). \tag{5.2.10}$$

KGF-i omadus (ühemõõtmelisel juhul):

$$K_x'(0) = \mathbb{E}\,x \tag{5.2.11}$$

$$K_x''(0) = \operatorname{Var} x \tag{5.2.12}$$

$$K_x'''(0) = \mathbb{E}(x - \mathbb{E}\,x)^3 \tag{5.2.13}$$

ja kahemõõtmelisel juhul:

$$\frac{\partial^2 K(0,0)}{\partial s_1 \partial s_2} = \operatorname{Cov}(x_1, x_2). \tag{5.2.14}$$

5.2.7 Probability Generating Function

For discrete, non-negative random variables

$$G(z) = \mathbb{E}(z^X) = \sum_{x=0}^{\infty} p(x) z^x. \tag{5.2.15}$$

Properties:

$$p(k) = \Pr(X = k) = \frac{G^{(k)}(0)}{k!} \tag{5.2.16}$$

5.2.8 Distribution Function of a Function of Random Variable

Olgu juhuslikud muutujad X ja Y kusjuures $X = X(Y)$. Siis

$$F_x(x) = \Pr[X < x] = \Pr[X < x(y)] = F_x[x(y)] = F_y(y) \tag{5.2.17}$$

ja

$$f_y(y) = F_y'(y) = \frac{\mathrm{d}}{\mathrm{d}y} F_x[x(y)] = \frac{\mathrm{d}}{\mathrm{d}x} F_x(x) \frac{\mathrm{d}x}{\mathrm{d}y} = f_x[x(y)] \frac{\mathrm{d}x}{\mathrm{d}y}. \tag{5.2.18}$$

5.3 Ühemõõtmelised diskreetsed jaotused

5.3.1 Bernoulli jaotus

Kõige lihtsam kulli-ja-kirja jaotus. Olgu sündmuse A tõenäosus p. Siis juhuslik muutuja

$$Y = \begin{cases} 1 & \text{kui } A, \\ 0 & \text{kui } \bar{A}. \end{cases} \tag{5.3.1}$$

on Bernoulli jaotusega, kusjuures $\mathbb{E}\, Y = p$ ja $\operatorname{Var} Y = p(1-p)$.

5.3.2 Binoomjaotus

On N ühesuguse sõltumatu Bernoulli jaotusega juhusliku muutuja summa jaotus. Olgu $X = \sum^N Y_i$ kus Y_i on Bernoulli jaotusega parameetriga p. Siis

$$\Pr(X = x) = \binom{N}{x} p^x (1-p)^{N-x} \quad x \in \{0, \ldots, N\} \tag{5.3.2}$$

$$EY = Np \tag{5.3.3}$$

$$E(Y - EY)^2 = Np(1-p) \tag{5.3.4}$$

$$E(Y - EY)^3 = Np(1-p)(1-2p) \tag{5.3.5}$$

$$E(Y - EY)^4 = Np(1-p)(-1+3p-3p^2) \tag{5.3.6}$$

Tõestus: ühe katse korral kirjuta lahti $(y-p)^n$, arvesta et $Ey^n = p$ $(i \neq 0)$. N sõltumatu katse korral momendid liituvad.

5.3.3 Diskreetne jaotus

Jaotus kus juhslikul muutujal võib olla lõplik hulk diskreetseid väärtusi.

5.3.4 Geomeetriline jaotus ($Geo(p)$)

Geomeetriline jaotus kirjeldab mingi hulga Bernoulli jaotusega suuruste järjest esinemist. Olgu sündmuse tõenäosus p. Tõenäosus, et järjest toimub n sündmust ja seejärel sündmuste jada katkeb on:

$$f(n) = (1-p)p^n. \tag{5.3.7}$$

Jaotuse omadused:

$$\mathbb{E}\, n = \frac{1-p}{p} \tag{5.3.8}$$

$$\operatorname{Var} n = \frac{1-p}{p^2} \tag{5.3.9}$$

5.3.5 Multinoomjaotus

Olgu üksikul katsel M võimalikku tulemust $A_1 \ldots A_M$ vastavate tõenäosustega $p_1 \ldots p_M$, kusjuures $\sum^M p_i = 1$. Olgu N-katselises seerias N_i realiseerunud sündmuste A_i arv. Siis:

$$EN_i = Np_i \tag{5.3.10}$$

$$\mathrm{Var}\, N_i = Np_i(1-p_i) \tag{5.3.11}$$
$$\dots$$
$$\mathrm{Cov}(N_i, N_j) = -Np_ip_j \tag{5.3.12}$$

Tõestus: momentide arvutamisel võib multinoomjaotuse taandada binoomjaotuseks, kovariatsiooni jaoks kirjuta definitsioon lahti, arvesta et $EA_iA_j = 0$.

5.3.6 Poissoni jaotus

Poissoni jaotusega on sõltumatute sündmuste arv ajaühikus. Kui ajaühikus toimub keskmiselt λ sündmust, siis tõenäosus, et toimub n sündmust aja t jooksul on:

$$f(n) \equiv p_p(n) = \frac{\mathrm{e}^{-\lambda t}(\lambda t)^n}{n!} \tag{5.3.13}$$
$$P_p(n) \equiv \sum_{s=1}^{n} p_p(s) \tag{5.3.14}$$

Jaotuse omadused: jaotus on log-kumer,

$$\mathbb{E}\, n = \lambda t \qquad \mathrm{Var}\, n = \lambda t \tag{5.3.15}$$

Tõenäosus, et mingi aja t jooksul ei toimu ühtegi sündmust, $f(0)$, on eksponentjaotusega $\mathcal{E}(\lambda)$.

ML-hinnang: Kui vaatluse i jooksul, mille kestus on t_i toimub n_i sündmust, siis kõigi vaatluste ML hinnang on:

$$\hat{\lambda} = \frac{\sum n_i}{\sum t_i} \qquad \text{ja} \qquad \mathrm{Var}\, \hat{\lambda} = \frac{\sum n_i}{(\sum t_i)^2} \tag{5.3.16}$$

Poissoni summa tuletis aja järgi: Olgu

$$\vartheta(t) = \sum_{s=0}^{S} Q(s) \frac{(\lambda t)^2}{s!} \mathrm{e}^{-\lambda t} = \mathop{\mathbb{E}}_{s} Q(s) \tag{5.3.17}$$

siis

$$\frac{\partial}{\partial t}\vartheta(t) = \lambda \sum_{s=0}^{S-1} [Q(s+1) - Q(s)]\, p_p(s) - \lambda Q(S) p_p(S) = \tag{5.3.18}$$
$$= \lambda \sum_{s=1}^{S} Q(s) \left[p_p(s-1) - p(s)\right] - \lambda Q(0) p_p(0) \tag{5.3.19}$$
$$\frac{\partial}{\partial t} P_p(s) = -\lambda p_p(s) \tag{5.3.20}$$
$$\frac{\partial}{\partial t} p_p(s) = \begin{cases} -\lambda p_p(s) & \text{kui } s = 0 \\ -\lambda p_p(s) + \lambda p_p(s-1) & \text{kui } s > 0 \end{cases} \tag{5.3.21}$$

5.3.7 Skellam Distribution $PD(\lambda, \delta)$

Skellam distribution is distribution of difference of two independent Poisson RV-s. Let $N = X - Y$ where $X \sim Poisson(\lambda)$ and $Y \sim Poisson(\delta)$:

$$P(N = n) = \mathrm{e}^{-(\lambda+\delta)} \left(\frac{\lambda}{\delta}\right)^{\frac{n}{2}} I_n(2\sqrt{\lambda\delta}) \tag{5.3.22}$$

$$\mathbb{E}\, N = \lambda - \delta \tag{5.3.23}$$

$$\operatorname{Var} N = \lambda + \delta \tag{5.3.24}$$

5.4 1D Continuous Distributions

5.4.1 Eksponentjaotus $\mathcal{E}(\theta)$

Eksponentjaotus kirjeldab konstantse kiirusega hääbuvaid protsesse. Tõenäosustihedus:

$$f(t) = \theta e^{-\theta t}, \qquad t \geq 0, \theta > 0 \tag{5.4.1}$$

jaotusfunktsioon:

$$F(t) = 1 - e^{-\theta t}, \tag{5.4.2}$$

momendifunktsioon:

$$M_T(s) = \frac{1}{1 - \frac{s}{\theta}}, \quad s < \theta. \tag{5.4.3}$$

Momendid:

$$\mathbb{E}\, T = \frac{1}{\theta} \tag{5.4.4}$$

$$\mathbb{E}\, T^2 = \frac{2}{\theta^2} \tag{5.4.5}$$

$$\mathrm{Var}\,(T - \mathbb{E}(T))^2 = \frac{1}{\theta^2}. \tag{5.4.6}$$

$\log T$ on esimest liiki ekstreemväärtuste jaotusega. $\log T$-ga seotud suurused on:

$$M_{\log T}(s) = \frac{\Gamma(s+1)}{\theta^s} \tag{5.4.7}$$

$$K_{\log T}(s) = \log \Gamma(s+1) - s \log \theta \tag{5.4.8}$$

$$\mathbb{E} \log T = \psi(1) - \log \theta \tag{5.4.9}$$

$$\mathrm{Var}\ \log T = \psi'(1), \tag{5.4.10}$$

kus ψ on digamma funktsioon.

Kui $z_1 \sim \mathcal{E}(\theta_1)$ ja $z_2 \sim \mathcal{E}(\theta_2)$ siis

$$\log z_1 - \log z_2 \sim \frac{\theta_1}{\theta_1 + \theta_2 \mathrm{e}^{-x}} \sim \Lambda(x), \qquad \text{kui} \quad \theta_1 = \theta_2 \tag{5.4.11}$$

Tõestus: arvesta et $\Pr(z_1/z_2 < \alpha) = \Pr(z_1 < \alpha z_2)$ ja integreeri.

5.4.2 F-jaotus $F(n_1, n_2)$

F-jaotus tekib kahe χ^2 jaotusega suuruse jagamisel. Kui $w_1 \sim \chi^2_{n_1}$ ja $w_2 \sim \chi^2_{n_2}$ siis

$$\frac{\frac{w_1}{n_1}}{\frac{w_2}{n_2}} \sim F(n_1, n_2). \tag{5.4.12}$$

Tihedusfunktsioon:

$$f(x) = \frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} x^{\frac{n_1}{2}-1}}{B\left(\frac{n_1}{2}, \frac{n_2}{2}\right)\left(1 + \frac{n_1}{n_2}x\right)^{\frac{n_1+n_2}{2}}} \tag{5.4.13}$$

5.4.3 Gammajaotus $\mathcal{G}(\alpha, \beta)$

Kasutatakse heterogeensuse kirjeldamiseks kestusmudelites. Tõenäosustihedus:

$$f(x) = \frac{1}{\beta^\alpha} \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \qquad x > 0, \tag{5.4.14}$$

$\alpha > 0$ kirjeldab kuju ja $\beta > 0$ on skaalaparameeter. Märkus: mõnikord kasutatakse ka parameetreid $(\alpha, \frac{1}{\beta})$.

Momendifunktsioon ja ootused:

$$\begin{aligned}
M_x(s) &= \frac{1}{(\beta s - 1)^\alpha} && (5.4.15)\\
K_x(s) &= -\alpha \log(\beta s - 1) && (5.4.16)\\
\mathbb{E}\, x &= \beta\alpha && (5.4.17)\\
\mathbb{E}\, x^2 &= \beta^2 \alpha(\alpha + 1) && (5.4.18)\\
\operatorname{Var} x &= \beta^2 \alpha. && (5.4.19)
\end{aligned}$$

$\log x$ kirjeldav momendifunktsioon ja ootused:

$$\begin{aligned}
M_{\log x}(s) &= \frac{\Gamma(s + \alpha)}{\Gamma(\alpha)} \beta^s && (5.4.20)\\
K_{\log x}(s) &= s \log \beta + \log \Gamma(s + \alpha) - \log \Gamma(\alpha) && (5.4.21)\\
\mathbb{E} \log x &= \log \beta + \psi(\alpha) && (5.4.22)\\
\operatorname{Var} \log x &= \psi'(\alpha) && (5.4.23)
\end{aligned}$$

Erijuht on *normaalne gammajaotus*, mille keskväärtus on 1. Sel juhul

$$\alpha = \frac{1}{\beta} \equiv \eta \tag{5.4.24}$$

ja jaotusfunktsioon

$$f_x(x) = \eta^\eta \frac{1}{\Gamma(\eta)} x^{\eta-1} \mathrm{e}^{-\eta x}. \tag{5.4.25}$$

Sel juhul:

$$\begin{aligned}
\mathbb{E}\, x &= 1 && (5.4.26)\\
\mathbb{E}\, x^2 &= 1 + \frac{1}{\eta} && (5.4.27)\\
\operatorname{Var} x &= \frac{1}{\eta} && (5.4.28)
\end{aligned}$$

Teine oluline erijuht on χ^2-jaotus. Kui $\alpha = \frac{k}{2}$ ja $\beta = 2$, siis X jaotusfunktsioon on

$$f_x(x) = \frac{1}{2^{\frac{k}{2}}} \frac{1}{\Gamma(\frac{k}{2})} y^{\frac{k}{2}-1} \mathrm{e}^{-\frac{y}{2}}. \tag{5.4.29}$$

Tolle kohta öeldakse $\chi^2(k)$ jaotus.

5.4.4 Hii-ruut jaotus $\chi^2(k)$

$\chi^2(k)$ jaotus tekib kui liita kokku k normaaljaotusega juhusliku suuruse ruutu. Jaotusfunktsioon:

$$f_x(x) = \frac{1}{2^{\frac{k}{2}}} \frac{1}{\Gamma(\frac{k}{2})} y^{\frac{k}{2}-1} \mathrm{e}^{-\frac{y}{2}}. \tag{5.4.30}$$

5.4.5 Log-normaalne jaotus $LN(\mu, \sigma^2)$

Log-normaalse jaotusega on juhuslik suurus, mille logaritm on normaaljaotusega. Tihedusfunktsioon:

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma x} \mathrm{e}^{-\frac{1}{2}\left[\frac{(\log x - \mu)}{\sigma}\right]^2}. \tag{5.4.31}$$

Omadused:

$$\mathbb{E}\, x = \mathrm{e}^{\mu + \frac{1}{2}\sigma^2} \tag{5.4.32}$$

$$\operatorname{Var} x = \mathrm{e}^{2\mu + \sigma^2}(\mathrm{e}^{\sigma^2} - 1). \tag{5.4.33}$$

5.4.6 Log-ühtlane jaotus

Kasutatakse palgajaotuse kirjeldamiseks. Tihedusfunktsioon:

$$f(x) = \frac{1}{x} \frac{1}{\log \beta - \log \alpha}, \qquad 0 \le \alpha \le x \le \beta < \infty. \tag{5.4.34}$$

5.4.7 Logistic Distribution

$$\text{cdf} \qquad \Lambda(x) = \frac{e^x}{1 + e^x} = \frac{1}{1 + \mathrm{e}^{-x}} \tag{5.4.35}$$

$$\text{pdf} \qquad f(x) = \frac{\mathrm{e}^x}{(1 + \mathrm{e}^x)^2} = \frac{\mathrm{e}^{-x}}{(1 + \mathrm{e}^{-x})^2} \tag{5.4.36}$$

$$f'(x) = \mathrm{e}^{-x} \frac{\mathrm{e}^{-x} - 1}{(\mathrm{e}^{-x} + 1)^3} \tag{5.4.37}$$

$$\text{MGF} \qquad M(s) = \int \frac{\mathrm{e}^{x(s+1)}}{(1 + \mathrm{e}^x)^2} \mathrm{d}x \tag{5.4.38}$$

Logistic distribution is symmetric around 0, i.e. $\Lambda(x) = 1 - \Lambda(-x)$ and $f(x) = f(-x)$.

5.4.8 Normal Distribution $N(\mu, \sigma^2)$

Sum of many independent random disturbations tends to be normally distributed (Central Limit Theorem.)

$$F(x; \mu, \sigma) \equiv \Phi\left(\frac{x - \mu}{\sigma}\right) \qquad \text{cannot be expressed analytically} \tag{5.4.39}$$

$$f(x; \mu, \sigma) \equiv \frac{1}{\sigma} \phi\left(\frac{x - \mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \mathrm{e}^{\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} \tag{5.4.40}$$

Characteristic function:

$$\phi_X(t) = e^{it\mu - \frac{1}{2}\sigma^2 t^2} \tag{5.4.41}$$

Moment generating function

$$M_x(s) = e^{\mu s + \frac{1}{2}\sigma^2 s^2} \tag{5.4.42}$$

Properties: if $X_i \sim N(\mu_i, \sigma_i^2)$ are independent normals

$$\sum_i X_i \sim N\left(\sum_i \mu_i, \sum_i \sigma_i^2\right). \tag{5.4.43}$$

Conditional Expectations If $X \sim N(\mu, \sigma)$,

$$\mathbb{E}[X|X < a] = \mu - \sigma \frac{\phi(\frac{a-\mu}{\sigma})}{\Phi(\frac{a-\mu}{\sigma})} = \mu - \sigma\lambda\left(\frac{a-\mu}{\sigma}\right) \tag{5.4.44}$$

$$\mathbb{E}[X|X > a] = \mu + \sigma \frac{\phi(\frac{a-\mu}{\sigma})}{1 - \Phi(\frac{a-\mu}{\sigma})} = \mu + \sigma\lambda\left(\frac{\mu - a}{\sigma}\right) \tag{5.4.45}$$

$$\mathbb{E}[X|X \in [a, b]] = \mu - \sigma \frac{\phi\left(\frac{b-\mu}{\sigma}\right) - \phi\left(\frac{a-\mu}{\sigma}\right)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)} \tag{5.4.46}$$

If $X \sim N(0, \sigma)$,

$$\mathbb{E}[X|X < a] = -\sigma\lambda(\frac{a}{\sigma}) \tag{5.4.47}$$

$$\mathbb{E}[X|X > a] = \sigma\lambda(-\frac{a}{\sigma}) \tag{5.4.48}$$

$$\mathbb{E}[X^2|X < a] = \sigma^2 - \sigma a\lambda\left(\frac{a}{\sigma}\right) \tag{5.4.49}$$

$$\mathbb{E}[X^2|X > a] = \sigma^2 + \sigma a\lambda\left(-\frac{a}{\sigma}\right) \tag{5.4.50}$$

$$\mathbb{E}[X^2|X > -a \wedge X < a] = \sigma^2 - 2\frac{\sigma a\phi\left(\frac{a}{\sigma}\right)}{1 - 2\Phi\left(\frac{a}{\sigma}\right)} \tag{5.4.51}$$

$$\mathbb{E}[X^2|X < -a \vee X > a] = \mathbb{E}[X^2|X > a] \tag{5.4.52}$$

$$\mathrm{Var}\,[X|X < a] = \sigma^2\left[1 - \frac{a}{\sigma}\lambda\left(\frac{a}{\sigma}\right) - \lambda^2(\frac{a}{\sigma})\right] \tag{5.4.53}$$

$$\mathrm{Var}\,[X|X > a] = \sigma^2\left[1 + \frac{a}{\sigma}\lambda\left(-\frac{a}{\sigma}\right) - \lambda^2(-\frac{a}{\sigma})\right] \tag{5.4.54}$$

$$\tag{5.4.55}$$

Let $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$, $X \perp\!\!\!\perp Y$. Now

$$\mathbb{E}[X|X < Y] = \mu_X - \frac{\sigma_X^2}{\sqrt{\sigma_X^2 + \sigma_Y^2}}\lambda\left(-\frac{\mu_X - \mu_Y}{\sqrt{\sigma_X^2 + \sigma_Y^2}}\right) \tag{5.4.56}$$

$$\mathbb{E}[X|X > Y] = \mu_X + \frac{\sigma_X^2}{\sqrt{\sigma_X^2 + \sigma_Y^2}} \lambda \left(\frac{\mu_X - \mu_Y}{\sqrt{\sigma_X^2 + \sigma_Y^2}} \right). \tag{5.4.57}$$

Proof: write $\mathbb{E}[X|X < Y] = \mathbb{E}[X|Z < 0]$ where $Z = X - Y$. Now follows from (5.5.9)

5.4.9 Pareto Distribution

Describes the upper part of many highly unequal distributions.

$$F_X(x) = 1 - \left(\frac{x_0}{x} \right)^{\alpha}, \quad x \geq x_0 > 0; \quad \alpha > 0 \tag{5.4.58}$$

$$f_X(x) = \alpha\, x_0^{\alpha}\, x^{-\alpha-1} \tag{5.4.59}$$

$$\mathbb{E}\, X = \frac{\alpha}{\alpha - 1} x_0, \quad \alpha > 1 \tag{5.4.60}$$

Properties:

- power law: $\log f_X(x) = log\left(\alpha\, x_0^{\alpha}\right) - (\alpha + 1) \log x$ is linear on log-log scale.
- it is *scale-free*: there is no features in the right tail, wherever you look, you have most observations that are smaller, but you also have observations that are way larger.

5.4.10 Pööratud normaaljaotus

Tihedusfunktsioon:

$$f(t) = \frac{1}{t^{\frac{3}{2}}} \phi \left(\frac{\mu t - 1}{\sigma \sqrt{t}} \right) \tag{5.4.61}$$

ja jaotusfunktsioon:

$$F(t) = \Phi \left(\frac{\mu t - 1}{\sigma \sqrt{t}} \right) - e^{2 \frac{\mu}{\sigma^2}} \Phi \left(- \frac{\mu t + 1}{\sigma \sqrt{t}} \right). \tag{5.4.62}$$

Kɵik momendid on olemas kui $\mu > 0$:

$$\mathbb{E}\, T = \frac{1}{\mu} \tag{5.4.63}$$

$$\operatorname{Var} T = \frac{\sigma^2}{\mu^3}. \tag{5.4.64}$$

Kui $\mu = 0$, on jaotus korralik, positiivsed momendid aga puuduvad.

5.4.11 t-Distribution

Used for *t-test*. For *n degrees of freedom*:

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{w}\right)} \left(1 + \frac{x^2}{n} \right)^{-\frac{n+1}{2}} \tag{5.4.65}$$

$$\mathbb{E}\,X = 0 \tag{5.4.66}$$
$$\operatorname{Var} X = \frac{n}{n-2} \tag{5.4.67}$$
$$\text{skewness } g_1 = 0 \tag{5.4.68}$$
$$\text{curtosis } g_2 = \frac{3n-6}{n-4} \qquad (n > 4) \tag{5.4.69}$$

5.4.12 Triangular Distribution

$$F(x) = \begin{cases} 0 & x < 0 \\ x^2 & 0 \le x \le 1 \\ 1 & x > 1 \end{cases} \tag{5.4.70}$$
$$f(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & \text{elsewhere} \end{cases} \tag{5.4.71}$$
$$\mathbb{E}\,X = \frac{2}{3} \tag{5.4.72}$$
$$\mathbb{E}\,X^2 = \frac{1}{2} \tag{5.4.73}$$
$$\operatorname{Var} X = \frac{1}{18} \tag{5.4.74}$$

5.4.13 Type-1 Extreme Value Distribution EV_1

(Also Gumbel distribution.) Describes $\log T$ when $T \sim \mathcal{E}(1)$.

$$f(x) = \mathrm{e}^{-x}\mathrm{e}^{-\mathrm{e}^{-x}} \tag{5.4.75}$$
$$F(x) = \mathrm{e}^{-\mathrm{e}^{-x}}. \tag{5.4.76}$$
$$\mathbb{E}\,x = -c \approx 0,5772 \tag{5.4.77}$$

5.4.14 Type-2 Extreme Value Distribution EV_2

(Also Fréchet distribution or inverse Weibull distribution.)

$$f(x) = \alpha x^{-1-\alpha}\mathrm{e}^{-x^{-\alpha}} \tag{5.4.78}$$
$$F(x) = \mathrm{e}^{-x^{-\alpha}} \tag{5.4.79}$$

5.4.15 Uniform Distribution $Unif(a, b)$

$$\text{expectation} \quad \mu \equiv \mathbb{E}\,X = \frac{1}{2}(a+b) \tag{5.4.80}$$
$$\text{variance} \quad \sigma^2 \equiv \mathbb{E}(X-\mu)^2 = \frac{1}{12}(b-a) \tag{5.4.81}$$
$$\text{4th central moment} \quad m_4 \equiv \mathbb{E}(X-\mu)^4 = \frac{(b-a)^4}{80} \tag{5.4.82}$$

5.4.16 Weibulli jaotus

Weibulli jaotus on eksponentjaotuse üldistus, kasutatakse ajas ühtlaselt kahaneva hasardi kirjeldamiseks. Omadused:

$$F(t) = 1 - e^{-(\lambda t)^{\alpha}} \tag{5.4.83}$$

$$f(t) = \alpha \lambda^{\alpha} t^{\alpha-1} e^{-(\lambda t)^{\alpha}} \tag{5.4.84}$$

$$\theta(t) = \alpha \lambda^{\alpha} t^{\alpha-1} \tag{5.4.85}$$

where λ is scale- and α is the shape parameter.

5.5 Multivariate Continuous Distributions

5.5.1 Multivariate Normal $N(\mu, \Sigma)$

N-mõõtmelise normaajaotuse jaotusfunktsioon: Olgu

$$\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \tag{5.5.1}$$

kus $\boldsymbol{\mu}$ on keskväärtus ja $\boldsymbol{\Sigma}$ dispersioonimaatriks. Siis:

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = (2\pi)^{-\frac{n}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \mathrm{e}^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}. \tag{5.5.2}$$

2D Conditional Distributions Let $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}\right)$.

The exponent in the distribution function can be expressed as

$$-\frac{1}{2}\left(\frac{\sigma_2 x_1^2 - 2\sigma_{12} x_1 x_2 + \sigma_1^2 x_2^2}{\sigma_1^2 \sigma_2^2 - \sigma_{12}^2}\right) = -\frac{1}{2}\left[\frac{\left(x_1 - \frac{\sigma_{12}}{\sigma_2^2} x_2\right)^2}{\sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}} + \frac{x_2^2}{\sigma_2^2}\right]. \tag{5.5.3}$$

Accordingly, based on Bayesian law the probability density of 2D normal $f_{X_1,X_2}(x_1, x_2) = f_{X_1|X_2}(x_1, x_2) f_{X_2}(x_2) = f_{X_2|X_1}(x_1, x_2) f_{X_1}(x_1)$ where all the conditional and marginal distribution functions are normal:

$$(X_1|X_2 = x_2) \sim N\left(\frac{\sigma_{12}}{\sigma_2^2} x_2, \sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}\right) \qquad X_2 \sim N(0, \sigma_2^2) \tag{5.5.4}$$

$$(X_2|X_1 = x_1) \sim N\left(\frac{\sigma_{12}}{\sigma_1^2} x_1, \sigma_2^2 - \frac{\sigma_{12}^2}{\sigma_1^2}\right) \qquad X_1 \sim N(0, \sigma_1^2) \tag{5.5.5}$$

Distribution for $(X_1|X_2 < a)$:

$$f_{X1|X2<a}(x_1) = \frac{1}{\sigma_1} \frac{\phi\left(\frac{x_1}{\sigma_1}\right)}{\Phi\left(\frac{a}{\sigma_2}\right)} \Phi\left(\frac{a - \frac{\sigma_{12}}{\sigma_1^2} x_1}{\sqrt{\sigma_2^2 - \frac{\sigma_{12}^2}{\sigma_1^2}}}\right) \tag{5.5.6}$$

Distribution for $(X_1|X_2 > a)$:

$$f_{X1|X2>a}(x_1) = \frac{1}{\sigma_1} \frac{\phi\left(\frac{x_1}{\sigma_1}\right)}{\Phi\left(-\frac{a}{\sigma_2}\right)} \Phi\left(-\frac{a - \frac{\sigma_{12}}{\sigma_1^2} x_1}{\sqrt{\sigma_2^2 - \frac{\sigma_{12}^2}{\sigma_1^2}}}\right) \tag{5.5.7}$$

Conditional Expectations Let $\boldsymbol{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$ and $\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$:

$$(X_1|X_2 = x_2) \sim N\left(\mu_1 + \frac{\sigma_{12}}{\sigma_2^2}(x_2 - \mu_2), \sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}\right) \tag{5.5.8}$$

(follows from 5.5.3) and

$$\mathbb{E}[X_1|X_2 < a] = \mu_1 - \frac{\sigma_{12}}{\sigma_2}\lambda\left(\frac{a-\mu_2}{\sigma_2}\right) \tag{5.5.9}$$

$$\mathbb{E}[X_1|X_2 > a] = \mu_1 + \frac{\sigma_{12}}{\sigma_2}\lambda\left(\frac{\mu_2-a}{\sigma_2}\right) \tag{5.5.10}$$

$$\mathbb{E}[X_1^2|X_2 < a] = \sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^3}a\lambda\left(\frac{a}{\sigma_2}\right) \tag{5.5.11}$$

$$\mathbb{E}[X_1^2|X_2 > a] = \sigma_1^2 + \frac{\sigma_{12}^2}{\sigma_2^3}a\lambda\left(-\frac{a}{\sigma_2}\right) \tag{5.5.12}$$

$$\mathbb{E}[X_1^2|X_2 \in \mathcal{A}] = \frac{\sigma_{12}^2}{\sigma_2^4}\,\mathbb{E}[X_2^2|X_2 \in \mathcal{A}] + \sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2} \tag{5.5.13}$$

$$\operatorname{Var}[X_1|X_2 < a] = \sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^3}a\lambda\left(\frac{a}{\sigma_2}\right) - \frac{\sigma_{12}^2}{\sigma_2^2}\lambda^2\left(\frac{a}{\sigma_2}\right) \tag{5.5.14}$$

$$\operatorname{Var}[X_1|X_2 > a] = \sigma_1^2 + \frac{\sigma_{12}^2}{\sigma_2^3}a\lambda\left(-\frac{a}{\sigma_2}\right) - \frac{\sigma_{12}^2}{\sigma_2^2}\lambda^2\left(-\frac{a}{\sigma_2}\right) \tag{5.5.15}$$

Proof: write (5.5.8) $\Rightarrow X_1 = \mu_1 + \frac{\sigma_{12}}{\sigma_2}(X_2 - \mu_2) + E$, where E and X_2 are independent (property of normal distribution). Find $X_1|X_2 \in \mathcal{A} = \frac{\sigma_{12}}{\sigma_2^2}\,\mathbb{E}[X_2|X_2 \in \mathcal{A}] + E$.

Multiplying normals

$$\frac{1}{\sigma_1}\phi\left(\frac{x-ay}{\sigma_1}\right)\frac{1}{\sigma_2}\phi\left(\frac{y-b}{\sigma_2}\right) = \frac{1}{\sigma_x}\phi\left(\frac{x-ab}{\sigma_x}\right)\frac{1}{\sigma_y}\phi\left(\frac{y-\mu_y}{\sigma_y}\right), \tag{5.5.16}$$

where

$$\sigma_x^2 = \sigma_1^2 + \sigma_2^2a^2 \qquad \sigma_y = \frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2a^2}}$$

$$\mu_y = \frac{\sigma_1^2b + \sigma_2^2ax}{\sigma_1^2+\sigma_2^2a^2}$$

The same in multi-dimensional case:

$$\begin{aligned}
\frac{1}{\sigma_1}\phi\left(\frac{x_1-y}{\sigma_1}\right)\frac{1}{\sigma_1}\phi\left(\frac{x_2-y}{\sigma_1}\right)\dots\frac{1}{\sigma_1}\phi\left(\frac{x_n-y}{\sigma_1}\right)\frac{1}{\sigma_2}\phi\left(\frac{y}{\sigma_2}\right) &= \\
= \prod_{i=1}^{n}\frac{1}{\sigma_1}\phi\left(\frac{x_i-y}{\sigma_1}\right)\frac{1}{\sigma_2}\phi\left(\frac{y}{\sigma_2}\right) &= \\
= \frac{1}{\sigma_x}\phi\left(\frac{x_1}{\sigma_x}\right)\frac{1}{\sigma_x}\phi\left(\frac{x_2}{\sigma_x}\right)\dots\frac{1}{\sigma_x}\phi\left(\frac{x_n}{\sigma_x}\right)\frac{1}{\sigma_y}\phi\left(\frac{y-\mu_y}{\sigma_y}\right) &= \\
= \prod_{i=1}^{n}\frac{1}{\sigma_x}\phi\left(\frac{x_i}{\sigma_x}\right)\frac{1}{\sigma_y}\phi\left(\frac{y-\mu_y}{\sigma_y}\right) &
\end{aligned} \tag{5.5.17}$$

where

$$\sigma_x^2 = \sigma_1^2\frac{\sigma_1^2+n\sigma_2^2}{\sigma_1^2+(n-1)\sigma_2^2} \qquad \sigma_y = \frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2+n\sigma_2^2}}$$

$$\mu_y = \frac{\sigma_2^2 \sum_{i=1}^n x_i}{\sigma_1^2 + n\sigma_2^2}$$

5.6 Jaotuste pered

5.6.1 Stabiilne pere

Mittenegatiivsesse stabiilsesse perre kuuluvad jaotused, mille momendifunktsioon on

$$M_x(s) = e^{-s^\alpha}, \qquad 0 < \alpha \leq 1. \tag{5.6.1}$$

Stabiilsel pere omadused:

1. kui juhusliku muutuja X_i jaotusfunktsioon on G_α mis kuulub stabiilsesse perre, siis juhusliku muutuja

   $$Y = n^{-\frac{1}{\alpha}} \sum_{i=1}^{N} X_i \tag{5.6.2}$$

   jaotusfunktsioon on kah G_α.

2. Momendifunktsiooni tuletis

   $$M_x'(s) = -\alpha s^{\alpha-1} M(s) \tag{5.6.3}$$

6 Estimators

6.1 M-Estimators

6.1.1 Variance

Let an estimator solve

$$H = \sum_i h_i(\hat{\boldsymbol{\theta}}) = 0. \tag{6.1.1}$$

From Taylor approximation

$$\sum_i h_i(\hat{\boldsymbol{\theta}}) = \sum_i h_i(\boldsymbol{\theta}_0) + \frac{\partial}{\partial \boldsymbol{\theta}} \sum_i h_i(\boldsymbol{\theta})\Big|_{\boldsymbol{\theta}_0} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = 0 \tag{6.1.2}$$

from where

$$\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 = -\left(\frac{\partial}{\partial \boldsymbol{\theta}} \sum_i h_i(\boldsymbol{\theta})\Big|_{\boldsymbol{\theta}_0} \right)^{-1} \sum_i h_i(\boldsymbol{\theta}_0). \tag{6.1.3}$$

The estimate for variance is

$$\operatorname{Var} \hat{\boldsymbol{\theta}} = \hat{\mathsf{A}}^{-1} \widehat{\operatorname{Var} H} \hat{\mathsf{A}}^{-1} \tag{6.1.4}$$

where

$$\hat{\mathsf{A}} = \frac{\partial}{\partial \boldsymbol{\theta}} \sum_i h_i(\boldsymbol{\theta})\Big|_{\boldsymbol{\theta}_0} \tag{6.1.5}$$

and $\widehat{\operatorname{Var} H}$ is an estimator for $\operatorname{Var} H$.

6.2 Maximum likelihood

6.2.1 Relationship to Kulback-Leibler distance

Let the random variables $X_1, X_2, \dots, X_n$ be *i.i.d.* distributed according to a distribution function $F(\cdot|\boldsymbol{\vartheta})$ and corresponding density function $f(\cdot|\boldsymbol{\vartheta})$. Let $f(\cdot|\boldsymbol{\vartheta})$ be specified fully parametrically with a finite unknown parameter vector $\boldsymbol{\vartheta}$. The *log-likelihood* function of the observed values $x_1, x_2, \dots, x_n$ is:

$$\ell(\boldsymbol{\vartheta}|x_1, x_2, \dots, x_n) = \frac{1}{n} \sum_{i=1}^{n} \log f(x_i|\boldsymbol{\vartheta}). \tag{6.2.1}$$

The *maximum likelihood* estimator of $\boldsymbol{\vartheta}$ is the value of $\boldsymbol{\vartheta}$ which maximises the log-likelihood function:

$$\hat{\boldsymbol{\vartheta}} = \arg\max_{\boldsymbol{\vartheta}} \ell(\boldsymbol{\vartheta}|x_1, x_2, \dots, x_n). \tag{6.2.2}$$

This can be written as

$$\hat{\boldsymbol{\vartheta}} = \arg\max_{\boldsymbol{\vartheta}} \int \log f(x|\boldsymbol{\vartheta}) \, \mathrm{d}F_n(x), \tag{6.2.3}$$

where $F_n(x)$ is the empirical distribution function:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(X_i \le x). \tag{6.2.4}$$

Further, we may write the estimator as

$$\hat{\boldsymbol{\vartheta}} = \arg\min_{\boldsymbol{\vartheta}} \left[\int \log f(x|\boldsymbol{\vartheta}_0)\, \mathrm{d}F_n(x) - \int \log f(x|\boldsymbol{\vartheta})\, \mathrm{d}F_n(x) \right], \tag{6.2.5}$$

where $f(\cdot|\boldsymbol{\vartheta}_0)$ is the true density function of X that does not depend on $\boldsymbol{\vartheta}$. Hence

$$\hat{\boldsymbol{\vartheta}} = \arg\min_{\boldsymbol{\vartheta}} \int \log \frac{f(x|\boldsymbol{\vartheta}_0)}{f(x|\boldsymbol{\vartheta})}\, \mathrm{d}F_n(x) = \arg\min_{\boldsymbol{\vartheta}} KL(f|\boldsymbol{\vartheta}_0, f|\boldsymbol{\vartheta}) \tag{6.2.6}$$

where $KL(f, g)$ is Kulback-Leibler distance of two probability density functions

$$KL(f, g) = \mathop{\mathbb{E}}_{f} \left[\log \frac{f(X)}{g(X)} \right] \tag{6.2.7}$$

and $\mathbb{E}_f$ means expectation over X that is distributed according to f.

6.2.2 Information matrix

Information matrix is defined as

$$I(\boldsymbol{\vartheta}) \equiv -\mathbb{E}\left[\frac{\partial^2 \ell(\vartheta)}{\partial \boldsymbol{\vartheta} \partial \boldsymbol{\vartheta}'} \right] = \mathbb{E}\left[\frac{\partial \ell(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}} \frac{\partial \ell(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}'} \right] \tag{6.2.8}$$

6.3 Entropy Distance

6.3.1 Entropy Distance

Lin (1991) defines *entropy distance:*

$$K(p_1, p_2) = \sum_x p_1(x) \log_2 \frac{p_1(x)}{\frac{1}{2}[p_1(x) + p_2(x)]} = 1 + \frac{1}{\log 2} \sum_x p_1(x) \log \frac{p_1(x)}{p_1(x) + p_2(x)} \tag{6.3.1}$$

Properties: $K(p_1, p_2) = 0$ if and only if $p_1 \equiv p_2$. Otherwise, $K > 0$.

7 Stochastic Processes

7.1 Autoregressive (AR) Processes

AR(1) process Juhuslik muutuja U järgib AR(1) protsessi kui U käesoleva perioodi realisatsioon on seotud eelmise perioodi omaga

$$u_t = \varrho u_{t-1} + \varepsilon_t \tag{7.1.1}$$

ning ε_t väärtused eri ajaperioodidel on sõltumatud. Et protsess oleks stabiilne peab ϱ väärtus jääma vahemikku $(-1, 1)$.

AR(2) protsess Juhuslik muutuja U järgib AR(2) protsessi kui U käesoleva perioodi realisatsioon on seotud kahe eelmise perioodi omaga

$$u_t = \varrho_1 u_{t-1} + \varrho_2 u_{t-2} + \varepsilon_t \tag{7.1.2}$$

ning ε_t väärtused eri ajaperioodidel on sõltumatud.

7.2 Hulkumine

Definitsioon: hulkumine (*random walk*) on statistiline protsess

$$z_{t+1} = z_t + \varepsilon_{t+1}. \tag{7.2.1}$$

7.2.1 Hulkumine vastu barjääri

Olgu $z_0 = 0$ ja $\varepsilon \sim N(0, 1)$ *i.i.d.* protsess. Siis z_2 jaotus tingimusel et z_1 es ületa barjääri α on

$$f(z_2|z_1 < \alpha) = \frac{1}{\sqrt{2}} \frac{\Phi\left(\sqrt{2}\alpha - \frac{1}{\sqrt{2}} z_2\right)}{\Phi(\alpha)} \phi\left(\frac{z_2}{\sqrt{2}}\right) \tag{7.2.2}$$

Tõestus: kirjuta $\phi(x)$ lahti ja integreeri.

8 Statistilised mudelid

8.1 Tobit-2 model

Definition:

$$y_{1i}^* = \boldsymbol{z}_i'\boldsymbol{\gamma} + u_{1i} \tag{8.1.1}$$
$$y_{2i}^* = \boldsymbol{x}_i'\boldsymbol{\beta} + u_{2i} \tag{8.1.2}$$
$$y_{1i} = \begin{cases} 1, & \text{if} \quad y_{1i}^* > 0 \\ 0, & y_{1i}^* \leq 0. \end{cases} \tag{8.1.3}$$
$$y_{2i} = \begin{cases} y_{2i}^*, & \text{if} \quad y_{1i}^* > 0 \\ 0, & y_{1i}^* \leq 0. \end{cases} \tag{8.1.4}$$

Assume

$$\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \varrho \\ \varrho & \sigma^2 \end{pmatrix}\right). \tag{8.1.5}$$

The Heckman two-step estimator in this case is as follows: $\boldsymbol{\gamma}$ can be consistently estimated with probit model. Further we may write:

$$\mathbb{E}[Y_2|Y_1 > 0, \boldsymbol{x}, \boldsymbol{z}] = \boldsymbol{x}'\beta + \mathbb{E}[U_2|U_1 > -\boldsymbol{z}'\gamma] = \boldsymbol{x}'\beta + \varrho\sigma\lambda(-\boldsymbol{z}'\gamma) \tag{8.1.6}$$
$$\operatorname{Var}[Y_2|Y_1 > 0, \boldsymbol{x}, \boldsymbol{z}] = \mathbb{E}[U_2|U_1 > -\boldsymbol{z}'\gamma] = \sigma^2 + \varrho^2\sigma^2[-\boldsymbol{z}'\gamma\lambda(\boldsymbol{z}'\gamma) - \lambda^2(\boldsymbol{z}'\gamma)] \tag{8.1.7}$$

where $\lambda(x) = \phi(x)/\Phi(x)$; $\Phi(\cdot)$ and $\phi(\cdot)$ are the normal cumulative distribution function and density function respectively. ϱ and σ can be estimated regressing y_{2i} on $\boldsymbol{x}_i$ and $\lambda(-\boldsymbol{z}'\gamma)$. From the coefficient of the latter, β_λ and the residual variance s^2, one can isolate ϱ and σ:

$$\hat{\sigma}^2 = s^2 + \beta_\lambda^2[\lambda^2(\boldsymbol{z}'\gamma) - \boldsymbol{z}'\gamma\lambda(\boldsymbol{z}'\gamma)] \tag{8.1.8}$$
$$\hat{\varrho} = \frac{\beta_\lambda}{\hat{\sigma}}. \tag{8.1.9}$$

Note that $\hat{\varrho}$ need not to be in $[-1, 1]$.

Denote:

$$r = \sqrt{1 - \varrho^2} \tag{8.1.10}$$
$$u_{2i} = y_{2i} - \boldsymbol{x}_i'\boldsymbol{\beta} \tag{8.1.11}$$
$$B_i = \frac{\boldsymbol{z}_i'\boldsymbol{\gamma} + \dfrac{\varrho}{\sigma}u_{2i}}{r} \tag{8.1.12}$$
$$C(B) = -\frac{\Phi(B)\phi(B)B + \phi(B)^2}{\Phi(B)^2} \tag{8.1.13}$$

The contribution of observation i to the log-likelihood:

$$\ell = \sum_{i:y_{1i} \leq 0} \log \Phi(-\boldsymbol{z}_i'\boldsymbol{\gamma}) + \tag{8.1.14}$$

$$+\sum_{i:y_{1i}>0}\left[\log\Phi(B_i)-\frac{1}{2}\log 2\pi-\log\sigma-\frac{1}{2}\frac{u_{2i}^2}{\sigma^2}\right]. \tag{8.1.15}$$

The gradient of the log-likelihood is:

$$\frac{\partial\ell}{\partial\boldsymbol{\gamma}}=\sum_{i:y_{1i}\le 0}-\lambda(\boldsymbol{z}_i'\boldsymbol{\gamma})\boldsymbol{z}_i+\sum_{i:y_{1i}>0}\lambda(B_i)\frac{\boldsymbol{z}_i}{r} \tag{8.1.16}$$

$$\frac{\partial\ell}{\partial\boldsymbol{\beta}}=\sum_{i:y_{1i}>0}\left[\frac{u_{2i}}{\sigma^2}-\lambda(B_i)\frac{\varrho}{\sigma}\frac{1}{r}\right]\boldsymbol{x}_i \tag{8.1.17}$$

$$\frac{\partial l}{\partial\sigma}=\sum_{i:y_{1i}>0}\left[\frac{u_{2i}^2}{\sigma^3}-\frac{1}{\sigma}-\lambda(B_i)\frac{\varrho}{\sigma^2}\frac{u_{2i}}{r}\right] \tag{8.1.18}$$

$$\frac{\partial\ell}{\partial\varrho}=\sum_{i:y_{1i}>0}\lambda(B_i)\frac{\frac{1}{\sigma}u_{2i}+\varrho\boldsymbol{z}_i'\boldsymbol{\gamma}}{r^3}. \tag{8.1.19}$$

Hessian components are

$$\frac{\partial^2\ell}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}'}=-\sum_{i:y_{1i}=0}C(-\boldsymbol{z}_i'\boldsymbol{\gamma})\boldsymbol{z}_i\boldsymbol{z}_i'+\sum_{i:y_{1i}=1}\frac{C(B)}{r}\boldsymbol{z}_i\boldsymbol{z}_i' \tag{8.1.20}$$

$$\frac{\partial^2\ell}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\beta}'}=-\sum_{i:y_{1i}=1}C(B)\frac{1}{\sigma}\frac{\varrho}{r}\boldsymbol{z}_i\boldsymbol{x}_i' \tag{8.1.21}$$

$$\frac{\partial^2\ell}{\partial\boldsymbol{\gamma}\partial\sigma}=-\sum_{i:y_{1i}=1}C(B)\frac{\varrho u_2}{\sigma^2 r^2}\boldsymbol{z}_i \tag{8.1.22}$$

$$\frac{\partial^2\ell}{\partial\boldsymbol{\gamma}\partial\varrho}=\sum_{i:y_{1i}=1}\left[C(B)\frac{\frac{u2}{\sigma}+\varrho\boldsymbol{z}_i'\boldsymbol{\gamma}}{r^4}+\lambda(B)\frac{\varrho}{r^3}\right]\boldsymbol{z}_i \tag{8.1.23}$$

$$\frac{\partial^2\ell}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}=\sum_{i:y_{1i}=1}\frac{1}{\sigma^2}\left[\frac{\varrho^2}{r^2}\mathrm{C}(B)-1\right]\boldsymbol{x}_i\boldsymbol{x}_i' \tag{8.1.24}$$

$$\frac{\partial^2\ell}{\partial\boldsymbol{\beta}\partial\sigma}=\sum_{i:y_{1i}=1}\left[C(B)\frac{\varrho^2}{\sigma^3}\frac{u_2}{r^2}+\frac{\varrho}{\sigma^2}\frac{\lambda(B)}{r}-2\frac{u_2}{\sigma^3}\right]\boldsymbol{x}_i \tag{8.1.25}$$

$$\frac{\partial^2\ell}{\partial\boldsymbol{\beta}\partial\varrho}=\sum_{i:y_{1i}=1}\left[-C(B)\frac{\frac{u_2}{\sigma}+\varrho\boldsymbol{z}_i'\boldsymbol{\gamma}}{r^4}\frac{\varrho}{\sigma}-\frac{\lambda(B)}{\sigma}\frac{1}{r^3}\right]\boldsymbol{x}_i \tag{8.1.26}$$

$$\frac{\partial^2\ell}{\partial\sigma_3^2}=\sum_{i:y_{1i}=1}\left[\frac{1}{\sigma^2}-3\frac{u_2^2}{\sigma^4}+2\lambda(B)\frac{u2}{r}\frac{\varrho}{\sigma^3}+\frac{\varrho^2}{\sigma^4}\frac{u_2^2}{r^2}\mathrm{C}(B)\right] \tag{8.1.27}$$

$$\frac{\partial^2\ell}{\partial\sigma\partial\varrho}=-\frac{1}{r^3}\sum_{i:y_{1i}=1}\frac{u_2}{\sigma^2}\left[C(B)\frac{\varrho\left(\frac{u_2}{\sigma}+\varrho\boldsymbol{z}_i'\boldsymbol{\gamma}\right)}{r}+\lambda(B)\right] \tag{8.1.28}$$

$$\frac{\partial^2 \ell}{\partial \varrho^2} = \sum_{i:y_{1i}=1} \left[C(B) \left(\frac{\frac{u_2}{\sigma} + \varrho z_i' \gamma}{r^3} \right)^2 + \lambda(B) \frac{z_i' \gamma (1 + 2\varrho^2) + 3\varrho \frac{u_2}{\sigma}}{r^5} \right] \tag{8.1.29}$$

8.2 Tobit-2 Model with Binary Outcome

The underlying latent model:

$$y_i^{S*} = \boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S + \epsilon_i^S \tag{8.2.1}$$

$$y_i^{O*} = \boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O + \epsilon_i^0 \tag{8.2.2}$$

$$y_i^S = \begin{cases} 1, & \text{if} \quad y_i^{S*} > 0 \\ 0, & \quad y_i^{S*} \le 0. \end{cases} \tag{8.2.3}$$

$$y_i^O = \begin{cases} \text{undetermined}, & \text{if} \quad y_i^S = 0 \quad \text{(case 1)} \\ 0, & \text{if} \quad y_i^{O*} \le 0 \quad \text{and} \quad y_i^S = 1 \quad \text{(case 2)} \\ 1, & \text{if} \quad y_i^{O*} > 0 \quad \text{and} \quad y_i^S = 1 \quad \text{(case 3)} \end{cases} \tag{8.2.4}$$

Assume

$$\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \varrho \\ \varrho & 1 \end{pmatrix} \right). \tag{8.2.5}$$

The log-likelihood function contains 3 components, corresponding to the cases in (8.2.4):

$$\ell = \sum_{i \in \text{case 1}} \log \Phi\left(-\boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S\right) \tag{8.2.6}$$

$$+ \sum_{i \in \text{case 2}} \log \left[1 - \Phi\left(-\boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S\right) - \bar{\Phi}_2 \left(\begin{pmatrix} -\boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S \\ -\boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O \end{pmatrix}, \begin{pmatrix} 1 & \varrho \\ \varrho & 1 \end{pmatrix} \right) \right] \tag{8.2.7}$$

$$+ \sum_{i \in \text{case 3}} \log \bar{\Phi}_2 \left(\begin{pmatrix} -\boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S \\ -\boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O \end{pmatrix}, \begin{pmatrix} 1 & \varrho \\ \varrho & 1 \end{pmatrix} \right), \tag{8.2.8}$$

where $\bar{\Phi}_2(\cdot, \cdot)$ is the upper tail probability of 2-dimensional normal distribution.

Denote by $\mathcal{L}_i$ the corresponding individual likelihood value. The score vector:

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{\beta}^S} \ell = & \sum_{i \in \text{case 1}} \frac{1}{\mathcal{L}_i} \phi\left(-\boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S\right) \boldsymbol{x}_i^S \\ & + \sum_{i \in \text{case 2}} \frac{1}{\mathcal{L}_i} \phi\left(\boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S\right) \bar{\Phi}\left(\frac{\boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O - \varrho \boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S}{\sqrt{1 - \varrho^2}} \right) \boldsymbol{x}_i^S \\ & + \sum_{i \in \text{case 3}} \frac{1}{\mathcal{L}_i} \phi\left(\boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S\right) \Phi\left(\frac{\boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O - \varrho \boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S}{\sqrt{1 - \varrho^2}} \right) \boldsymbol{x}_i^S \\ \frac{\partial}{\partial \boldsymbol{\beta}^O} \ell = & \sum_{i \in \text{case 2}} \frac{1}{\mathcal{L}_i} \phi\left(\boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O\right) \bar{\Phi}\left(\frac{\boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S - \varrho \boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O}{\sqrt{1 - \varrho^2}} \right) \boldsymbol{x}_i^O \\ & + \sum_{i \in \text{case 3}} \frac{1}{\mathcal{L}_i} \phi\left(\boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O\right) \Phi\left(\frac{\boldsymbol{\beta}^{S\prime} \boldsymbol{x}_i^S - \varrho \boldsymbol{\beta}^{O\prime} \boldsymbol{x}_i^O}{\sqrt{1 - \varrho^2}} \right) \boldsymbol{x}_i^O \end{aligned}$$

$$\frac{\partial}{\partial\varrho}\ell = -\sum_{i\in\text{case 2}} \phi_2\left(\begin{pmatrix} -\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S \\ -\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O \end{pmatrix}, \begin{pmatrix} 1 & \varrho \\ \varrho & 1 \end{pmatrix}\right) + \sum_{i\in\text{case 3}} \phi_2\left(\begin{pmatrix} -\boldsymbol{\beta}^{S\prime}\boldsymbol{x}_i^S \\ -\boldsymbol{\beta}^{O\prime}\boldsymbol{x}_i^O \end{pmatrix}, \begin{pmatrix} 1 & \varrho \\ \varrho & 1 \end{pmatrix}\right),$$

where $\phi_2(\cdot,\cdot)$ is 2-dimensional normal density.

8.3 Tobit-5 Model

Definitsioon (lühiduse mõttes on indeks i ära jäetud):

$$
\begin{aligned}
y_1^* &= \boldsymbol{Z}'\boldsymbol{\gamma} + u_1 && (8.3.1)\\
y_2^* &= \boldsymbol{X}'\boldsymbol{\beta_2} + u_2 && (8.3.2)\\
y_3^* &= \boldsymbol{X}'\boldsymbol{\beta_3} + u_3 && (8.3.3)\\
y_2 &= \begin{cases} y_2^* & \text{kui} \quad y_1^* \le 0 \\ 0 & \phantom{\text{kui}} \quad y_1^* > 0 \end{cases} && (8.3.4)\\
y_3 &= \begin{cases} y_3^* & \text{kui} \quad y_1^* > 0 \\ 0 & \phantom{\text{kui}} \quad y_1^* \le 0 \end{cases} && (8.3.5)
\end{aligned}
$$

Eeldatakse et jääkliikmete jaotus on niisugune:

$$
\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \varrho_2\sigma_2 & \varrho_3\sigma_3 \\ \varrho_2\sigma_2 & \sigma_2^2 & \sigma_{23} \\ \varrho_3\sigma_3 & \sigma_{23} & \sigma_3^2 \end{pmatrix} \right). \tag{8.3.6}
$$

8.3.1 Heckmani kahesammuline hinnang

$\hat{\gamma}$ leitakse probiti abil. Edasi võib kirjutada

$$
\begin{aligned}
y_2 &= \boldsymbol{X}'\boldsymbol{\beta_2} - \varrho_2\sigma_2\lambda(-\boldsymbol{Z}'\boldsymbol{\gamma}) + e_2 \\
y_3 &= \boldsymbol{X}'\boldsymbol{\beta_3} + \varrho_3\sigma_3\lambda(\boldsymbol{Z}'\boldsymbol{\gamma}) + e_3
\end{aligned} \tag{8.3.7}
$$

Kusjuures

$$
\begin{aligned}
\sigma_{e2} &= \sigma_2^2\left\{1 - \varrho_2^2\left[\lambda^2(-\boldsymbol{Z}'\boldsymbol{\gamma}) - \boldsymbol{Z}'\boldsymbol{\gamma}\lambda(-\boldsymbol{Z}'\boldsymbol{\gamma})\right]\right\} \\
\sigma_{e3} &= \sigma_3^2\left\{1 - \varrho_3^2\left[\lambda^2(\boldsymbol{Z}'\boldsymbol{\gamma}) + \boldsymbol{Z}'\boldsymbol{\gamma}\lambda(-\boldsymbol{Z}'\boldsymbol{\gamma})\right]\right\}
\end{aligned} \tag{8.3.8}
$$

Kui lähendada seost (8.3.7) OLS-ga, siis saab λ koefitsendi ja dispersiooni hinnangu abil leida $\hat{\varrho}$ ja $\hat{\sigma}$. Märkus: $\hat{\varrho}$ ei pruugi olla -1 ja 1 vahel.

8.3.2 Maksimum-laiklikhuud hinnang

Mudeli log-laiklihuud on:

$$
\begin{aligned}
l &= -\frac{N}{2}\log 2\pi + \\
&+ \sum_{i \in \text{case } 2}\left\{-\log\sigma_2 - \frac{1}{2}\left(\frac{u_2}{\sigma_2}\right)^2 + \log\Phi\left[-\frac{\boldsymbol{Z}'\boldsymbol{\gamma} + \frac{\varrho_2}{\sigma_2}\left(y_2 - \boldsymbol{X}_i'\boldsymbol{\beta}_2\right)}{\sqrt{1-\varrho_2^2}}\right]\right\} \\
&+ \sum_{i \in \text{case } 3}\left\{-\log\sigma_3 - \frac{1}{2}\left(\frac{y_3 - \boldsymbol{X}\boldsymbol{\beta}_3}{\sigma_3}\right)^2 + \log\Phi\left[\frac{\boldsymbol{Z}'\boldsymbol{\gamma} + \frac{\varrho_3}{\sigma_3}\left(y_3 - \boldsymbol{X}_i'\boldsymbol{\beta}_3\right)}{\sqrt{1-\varrho_3^2}}\right]\right\}.
\end{aligned} \tag{8.3.9}
$$

Valikud 2 ja 3 erinevad ainult avaldise märgi poolest funktsiooni Φ sees. Tuletised on:

$$\frac{\partial l}{\partial \gamma} = -\sum_2 \frac{\phi(B_2)}{\Phi(B_2)} \frac{\boldsymbol{Z}}{\sqrt{1-\varrho_2^2}} + \sum_3 \frac{\phi(B_3)}{\Phi(B_3)} \frac{\boldsymbol{Z}}{\sqrt{1-\varrho_3^2}} \tag{8.3.10}$$

$$\frac{\partial l}{\partial \boldsymbol{\beta}_2} = \sum_2 \left[\frac{\phi(B_2)}{\Phi(B_2)} \left(\frac{\varrho_2}{\sigma_2} \frac{\boldsymbol{X}}{\sqrt{1-\varrho_2^2}} \right) + \frac{u_2}{\sigma_2^2} \boldsymbol{X} \right] \tag{8.3.11}$$

$$\frac{\partial l}{\partial \sigma_2} = \sum_2 \left[-\frac{1}{\sigma_2} + \frac{\left(y_2 - \boldsymbol{X}'\boldsymbol{\beta}_2\right)^2}{\sigma_2^3} + \frac{\phi(B_2)}{\Phi(B_2)} \frac{\varrho_2}{\sigma_2^2} \frac{y_2 - \boldsymbol{X}'\boldsymbol{\beta}_2}{\sqrt{1-\varrho_2^2}} \right] \tag{8.3.12}$$

$$\frac{\partial l}{\partial \varrho_2} = -\sum_2 \frac{\phi(B_2)}{\Phi(B_2)} \frac{\frac{1}{\sigma_2}(y_2 - \boldsymbol{X}'\boldsymbol{\beta}_2) + \varrho_2 \boldsymbol{Z}'\boldsymbol{\gamma}}{(1-\varrho_2^2)^{\frac{3}{2}}} \tag{8.3.13}$$

$$\frac{\partial l}{\partial \boldsymbol{\beta}_3} = \sum_3 \left[-\frac{\phi(B_3)}{\Phi(B_3)} \left(\frac{\varrho_3}{\sigma_3} \frac{\boldsymbol{X}}{\sqrt{1-\varrho_3^2}} \right) + \frac{u_3}{\sigma_3^2} \boldsymbol{X} \right] \tag{8.3.14}$$

$$\frac{\partial l}{\partial \sigma_3} = \sum_3 \left[-\frac{1}{\sigma_3} + \frac{\left(y_3 - \boldsymbol{X}'\boldsymbol{\beta}_3\right)^2}{\sigma_3^3} - \frac{\phi(B_3)}{\Phi(B_3)} \frac{\varrho_3}{\sigma_3^2} \frac{y_3 - \boldsymbol{X}'\boldsymbol{\beta}_3}{\sqrt{1-\varrho_3^2}} \right] \tag{8.3.15}$$

$$\frac{\partial l}{\partial \varrho_3} = \sum_3 \frac{\phi(B_3)}{\Phi(B_3)} \frac{\frac{1}{\sigma_3}(y_3 - \boldsymbol{X}'\boldsymbol{\beta}_3) + \varrho_3 \boldsymbol{Z}'\boldsymbol{\gamma}}{(1-\varrho_3^2)^{\frac{3}{2}}} \tag{8.3.16}$$

Teised tuletised:

$$\frac{\partial^2 l}{\partial \boldsymbol{\gamma}^2} = \sum_2 \frac{C(B_2)}{1-\varrho_2^2} z_i z_i' + \sum_3 \frac{C(B_3)}{1-\varrho_3^2} z_i z_i' \tag{8.3.17}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\beta}_2'} = -\sum_2 C(B_2) \frac{1}{\sigma_2} \frac{\varrho_2}{1-\varrho_2^2} \boldsymbol{Z}\boldsymbol{X}' \tag{8.3.18}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\gamma} \partial \sigma_2} = -\sum_2 \frac{\varrho_2 u_2}{\sigma_2^2 (1-\varrho_2^2)} C(B_2) \boldsymbol{Z} \tag{8.3.19}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\gamma} \partial \varrho_2} = \sum_2 \left[C(B_2) \frac{\frac{u_2}{\sigma_2} \varrho_2 \boldsymbol{Z}'\boldsymbol{\gamma}}{(1-\varrho_2^2)^2} - \lambda(B_2) \frac{\varrho_2}{(1-\varrho_2^2)^{\frac{3}{2}}} \right] \boldsymbol{Z} \tag{8.3.20}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\beta}_3'} = -\sum_3 C(B_3) \frac{1}{\sigma_3} \frac{\varrho_3}{1-\varrho_3^2} \boldsymbol{Z}\boldsymbol{X}' \tag{8.3.21}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\gamma} \partial \sigma_3} = -\sum_3 \frac{\varrho_3 u_3}{\sigma_3^2 (1-\varrho_3^2)} C(B_3) \boldsymbol{Z} \tag{8.3.22}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\gamma} \partial \varrho_3} = \sum_3 \left[C(B_3) \frac{\frac{u_3}{\sigma_3} \varrho_3 \boldsymbol{Z}'\boldsymbol{\gamma}}{(1-\varrho_3^2)^2} + \lambda(B_3) \frac{\varrho_3}{(1-\varrho_3^2)^{\frac{3}{2}}} \right] \boldsymbol{Z} \tag{8.3.23}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\beta}_2 \partial \boldsymbol{\beta}_2'} = \sum_2 \frac{1}{\sigma_2^2} \left[\frac{\varrho_2^2}{1-\varrho_2^2} C(B_2) - 1 \right] \boldsymbol{X}\boldsymbol{X}' \tag{8.3.24}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\beta}_2 \partial \sigma_2} = \sum_2 \left[C(B_2) \frac{u_2}{\sigma_2^3} \frac{\varrho_2^2}{1-\varrho_2^2} - \frac{\lambda(B_2)}{\sigma_2^2} \frac{\varrho_2}{\sqrt{1-\varrho_2^2}} - 2\frac{u_2}{\sigma_2^3} \right] \boldsymbol{X} \tag{8.3.25}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\beta}_2 \partial \varrho_2} = \sum_2 \left[-C(B_2) \frac{\frac{u_2}{\sigma_2} + \varrho_2 \boldsymbol{Z}'\boldsymbol{\gamma}}{(1-\varrho_2^2)^2} \frac{\varrho_2}{\sigma_2} + \frac{\lambda(B_2)}{\sigma_2} \frac{1}{(1-\varrho_2^2)^{\frac{3}{2}}} \right] \boldsymbol{X} \tag{8.3.26}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\beta}_2 \partial \boldsymbol{\beta}_3} = 0 \tag{8.3.27}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\beta}_2 \partial \sigma_3} = 0 \tag{8.3.28}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\beta}_2 \partial \varrho_3} = 0 \tag{8.3.29}$$

$$\begin{aligned} \frac{\partial^2 l}{\partial \sigma_2^2} &= \sum_2 \left[\frac{1}{\sigma_2^2} - 3\frac{u_2^2}{\sigma_2^4} + \frac{u_2}{\sigma_2^4} \frac{\varrho_2^2}{1-\varrho_2^2} C(B_2) \right] - \\ &- 2\sum_2 \lambda(B_2) \frac{u_2}{\sigma_2^3} \frac{\varrho_2}{\sqrt{1-\varrho_2^2}} \end{aligned} \tag{8.3.30}$$

$$\frac{\partial^2 l}{\partial \sigma_2 \partial \varrho_2} = \frac{1}{(1-\varrho_2^2)^{\frac{3}{2}}} \sum_2 \frac{u_2}{\sigma_2^2} \left[-C(B_2) \frac{\varrho_2 \left(\frac{u_2}{\sigma_2} + \varrho_2 \boldsymbol{Z}'\boldsymbol{\gamma} \right)}{\sqrt{1-\varrho_2^2}} + \lambda(B_2) \right] \tag{8.3.31}$$

$$\frac{\partial^2 l}{\partial \sigma_2 \partial \boldsymbol{\beta}_3} = 0 \tag{8.3.32}$$

$$\frac{\partial^2 l}{\partial \sigma_2 \partial \sigma_3} = 0 \tag{8.3.33}$$

$$\frac{\partial^2 l}{\partial \sigma_2 \partial \varrho_3} = 0 \tag{8.3.34}$$

$$\begin{aligned} \frac{\partial^2 l}{\partial \varrho_2^2} &= \sum_2 C(B_2) \left[\frac{\frac{u_2}{\sigma_2} + \varrho_2 \boldsymbol{Z}'\boldsymbol{\gamma}}{(1-\varrho_2^2)^{\frac{3}{2}}} \right]^2 - \\ &\quad - \sum_2 \frac{\phi(B_2)}{\Phi(B_2)} \frac{\boldsymbol{Z}'\boldsymbol{\gamma}(1+2\varrho_2^2) + 3\varrho_2 \frac{u_2}{\sigma_2}}{(1-\varrho_2^2)^{\frac{5}{2}}} \end{aligned} \tag{8.3.35}$$

$$\frac{\partial^2 l}{\partial \varrho_2 \partial \boldsymbol{\beta}_3} = 0 \tag{8.3.36}$$

$$\frac{\partial^2 l}{\partial \varrho_2 \partial \sigma_3} = 0 \tag{8.3.37}$$

$$\frac{\partial^2 l}{\partial \varrho_2 \partial \varrho_3} = 0 \tag{8.3.38}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\beta}_3 \partial \boldsymbol{\beta}_3'} = \sum_3 \frac{1}{\sigma_3^2}\left[\frac{\varrho_3^2}{1-\varrho_3^2}C(B_3) - 1\right]\boldsymbol{X}\boldsymbol{X}' \tag{8.3.39}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\beta}_3 \partial \sigma_3} = \sum_3 \left[C(B_3)\frac{\varrho_3^2}{\sigma_3^3}\frac{u_3}{1-\varrho_3^2} + \frac{\varrho_3}{\sigma_3^2}\frac{\lambda(B_3)}{\sqrt{1-\varrho_3^2}} - 2\frac{u_3}{\sigma_3^3}\right]\boldsymbol{X} \tag{8.3.40}$$

$$\frac{\partial^2 l}{\partial \boldsymbol{\beta}_3 \partial \varrho_3} = \sum_3 \left[-C(B_3)\frac{\frac{u_3}{\sigma_3} + \varrho_3 \boldsymbol{Z}'\boldsymbol{\gamma}}{(1-\varrho_3^2)^2}\frac{\varrho_3}{\sigma_3} - \frac{\lambda(B_3)}{\sigma_3}\frac{1}{(1-\varrho_3^2)^{\frac{3}{2}}}\right]\boldsymbol{X} \tag{8.3.41}$$

$$\begin{aligned}\frac{\partial^2 l}{\partial \sigma_3^2} &= \sum_3 \left[\frac{1}{\sigma_3^2} - 3\frac{u_3^2}{\sigma_3^4} + 2\lambda(B_3)\frac{y_3 - \boldsymbol{X}'\boldsymbol{\beta_3}}{\sqrt{1-\varrho_3^2}}\frac{\varrho_3}{\sigma_3^3}\right] + \\ &+ \sum_3 \frac{\varrho_3^2}{\sigma_3^4}\frac{u_3^2}{1-\varrho_3^2}C(B_3)\end{aligned} \tag{8.3.42}$$

$$\frac{\partial^2 l}{\partial \sigma_3 \partial \varrho_3} = -\frac{1}{(1-\varrho_3^2)^{\frac{3}{2}}}\sum_3 \frac{u_3}{\sigma_3^2}\left[C(B_3)\frac{\varrho_3\left(\frac{u_3}{\sigma_3} + \varrho_3 \boldsymbol{Z}'\boldsymbol{\gamma}\right)}{\sqrt{1-\varrho_3^2}} + \lambda(B_3)\right] \tag{8.3.43}$$

$$\begin{aligned}\frac{\partial^2 l}{\partial \varrho_3^2} &= \sum_3 C(B_3)\left[\frac{\frac{1}{\sigma_3}u_3 + \varrho_3 \boldsymbol{Z}'\boldsymbol{\gamma}}{(1-\varrho_3^2)^{\frac{3}{2}}}\right]^2 + \\ &+ \sum_3 \lambda(B_3)\frac{\boldsymbol{Z}'\boldsymbol{\gamma}(1+2\varrho_3^2) + 3\varrho_3\frac{1}{\sigma_3}u_3}{(1-\varrho_3^2)^{\frac{7}{2}}}\end{aligned} \tag{8.3.44}$$

Siin on tähistatud

$$B_2 = -\frac{\boldsymbol{Z}'\boldsymbol{\gamma} + \frac{\varrho_2}{\sigma_2}(y_2 - \boldsymbol{X}'\boldsymbol{\beta_2})}{\sqrt{1-\varrho_2^2}} \tag{8.3.45}$$

$$B_3 = \frac{\boldsymbol{Z}'\boldsymbol{\gamma} + \frac{\varrho_3}{\sigma_3}(y_3 - \boldsymbol{X}'\boldsymbol{\beta_3})}{\sqrt{1-\varrho_3^2}} \tag{8.3.46}$$

$$\lambda(B) = \frac{\phi(B)}{\Phi(B)} \tag{8.3.47}$$

$$u_2 = y_2 - \boldsymbol{X}'\boldsymbol{\beta_2} \tag{8.3.48}$$

$$u_3 = y_3 - \boldsymbol{X}'\boldsymbol{\beta_3} \tag{8.3.49}$$

$$C(B) = -\frac{\Phi(B)\phi(B)B + \phi(B)^2}{\Phi(B)^2} \tag{8.3.50}$$

8.4 Kestusmudelid

Tähistused:

τ kestus, algseisundis viibitud aeg

t kalendriaeg

8.4.1 Kaplan-Meieri hinnang

KM hinnang diskreetses ajas Olgu perioodil j r_j inimest "riski hulgas", s.t. r_j inimest võiksid põhimõtteliselt seisundist lahkuda. Lahkugu tegelikult n_j inimest, $r_j - n_j$ jäävad edasi algseisundisse. KM hinnang hasardile on seega

$$\hat{h} = \frac{n_j}{r_j} \tag{8.4.1}$$

$$\widehat{\operatorname{Var} \hat{h}_j} = \frac{\hat{h}_j(1 - \hat{h}_j)}{r_j}. \tag{8.4.2}$$

Kui periood j on k kuu pikkune, siis (keskmise) ühe kuu spetsiifilise hasardi saab

$$\hat{\vartheta}_j = 1 - (1 - \hat{h}_j)^{1/k} \tag{8.4.3}$$

$$\widehat{\operatorname{Var} \hat{\vartheta}_j} = \frac{\operatorname{Var} \hat{h}_j}{\left[k(1 - \hat{h}_j)^{1-1/k}\right]^2} \tag{8.4.4}$$

KM hinnang pidevas ajas Lahkugu aja t jooksul r algseisundis olnud inimesest n. Keskmine hasart ajaühikus on

$$\hat{\vartheta} = -\frac{1}{t} \log(1 - n/r) \tag{8.4.5}$$

$$\widehat{\operatorname{Var} \hat{\vartheta}} = \frac{1}{t^2} \frac{n/r}{r - n}. \tag{8.4.6}$$

8.4.2 Multiplikatiivne mittevaadeldav heterogeensus

Eeldame et hasart avaldub

$$\vartheta(\tau|x, v) = \lambda(\tau|x)v \tag{8.4.7}$$

kus v on mingi kindla jaotusega mittevaadeldav juhuslik suurus. Nüüd v keskväärtus algseisundisse jääjatel sõltub ajast:

$$\mathbb{E}(v|T \geq \tau) = -\frac{\mathcal{L}'[z(\tau|x)]}{\mathcal{L}[z(\tau|x)]} \tag{8.4.8}$$

kus on integreeritud hasart v-d arvestamata:

$$z(\tau|x) = \int_0^\tau \lambda(s|x)\,\mathrm{d}s \tag{8.4.9}$$

Kui v on algseisundisse sissevoolus ühikdispersiooniga gammajaotus parameetriga α, siis

$$\mathbb{E}(v|T \geq \tau) = \frac{\alpha}{z(\tau|x) + \alpha^{1/2}}. \tag{8.4.10}$$

8.4.3 Tükati konstantne põhihasart ja diskreetne mittevaadeldav heterogeensus ning pidev aeg

Sõltumatud vaatlused Eeldatakse, et hasart on konstantne iga M ajavahemiku sees, erinevatel ajavahemikel võib ta aga olla erinev. Olgu hasart kirjeldatud vektoriga $\boldsymbol{\lambda}$, kusjuures ajavahemiku j põhihasart olgu e^{λ_j}. Mittevaadeldav heterogeensus on diskreetse jaotusega:

$$v = \begin{cases} v_h \equiv 1, & \text{tõenäosusega} \quad p_h, \\ v_l, & \text{tõenäosusega} \quad p_l = 1 - p_h. \end{cases} \tag{8.4.11}$$

Sobiv on parametriseerida

$$\begin{array}{rclcrcl} v_1 & = & \mathrm{e}^{\tilde{v}_1} & & p_1 & = & \Lambda(\tilde{p}_1) \\ \dots & & & \text{ja} & \dots & & \\ v_K & = & 1 & & p_K & = & 1 - \sum^{K-1} p_k, \end{array} \tag{8.4.12}$$

Kus $\Lambda(\cdot)$ on logistile jaotusfunktsioon ja $v_k \in \Re$ ning $p_k \in \Re$.

Olgu m_i ajaperiood, mille jooksul inimene lahkub uuritavast seisundist ja tsenseerimist kirjeldagu $\delta_i = 0$ kui vaatlus on tsenseeritud ja 1 kui tsenseerimata ning $\mu_i = \mathrm{e}^{\gamma x_i'}$ olgu hasardi inimesest sõltuv osa. T_{ij} olgu teadaolev (võimalik et tsenseeritud) aeg, mis inimene i veetis uuritavas seisundis ajaperioodi j jooksul. Vektor $\boldsymbol{T}_i$ on M-vektor, mille komponendid on T_{ij} ning vector $\boldsymbol{d}_i$ on vektor, mille j-s komponent on 1, kui inimene lahkus uuritavast seisundist ajaperioodil j. Muud komponendid on nullid.

Sel juhul inimese i laiklihuud avaldub:

$$\mathcal{L}_i \quad = \quad p_l \mathcal{L}_{li} \; + \; p_h \mathcal{L}_{hi} \quad = \quad p_l \left(v_l \mu_i \mathrm{e}^{\lambda_{m_i}} \right)^{\delta_i} \mathrm{e}^{-z_{li}} \; + \; p_h \left(\mu_i \mathrm{e}^{\lambda_{m_i}} \right)^{\delta_i} \mathrm{e}^{-z_{hi}}, \tag{8.4.13}$$

kus

$$z_{li} = v_l \mu_i \sum_{j=1}^{M-1} \mathrm{e}^{\lambda_j} T_{ij} \qquad \text{ja} \qquad z_{hi} = \mu_i \sum_{j=1}^{M-1} \mathrm{e}^{\lambda_j} T_{ij} \tag{8.4.14}$$

on integreeritud hasart. Log-laiklihuudi gradient avaldub:

$$\frac{\partial \ell_i}{\partial v_l} = \frac{p_l \mathcal{L}_{li}}{\mathcal{L}} \left[\frac{\delta_i}{v_l} - z_{hi} \right] \tag{8.4.15}$$

$$\frac{\partial \ell_i}{\partial p_l} = \frac{1}{\mathcal{L}_i} \left[\mathcal{L}_{li} - \mathcal{L}_{hi} \right] \tag{8.4.16}$$

$$\frac{\partial \ell_i}{\partial \boldsymbol{\lambda}} = p_l \frac{\mathcal{L}_{li}}{\mathcal{L}_i} \left(\delta_i \boldsymbol{d}_i - v_l \mu_i \boldsymbol{T}_i \right) + p_h \frac{\mathcal{L}_{hi}}{\mathcal{L}_i} \left(\delta_i \boldsymbol{d}_i - \mu_i \boldsymbol{T}_i \right) \tag{8.4.17}$$

$$\frac{\partial \ell_i}{\partial \boldsymbol{\gamma}} = \frac{\boldsymbol{x_i}}{\mathcal{L}_i} \left[p_l \mathcal{L}_l \left(\delta_i - z_{li} \right) + p_h \mathcal{L}_h \left(\delta_i - z_{hi} \right) \right] \tag{8.4.18}$$

(8.4.19)

ja hessi maatriks:

$$\frac{\partial^2 \ell_i}{\partial v_l^2} = \frac{\mathcal{L}_{li}}{\mathcal{L}_i} \left[\left(\frac{\delta_i}{v_l} - z_{hi} \right)^2 \left(1 - p_l \frac{\mathcal{L}_{li}}{\mathcal{L}_i} \right) - \frac{\delta_i}{v_l^2} \right] \tag{8.4.20}$$

$$\frac{\partial^2 \ell_i}{\partial p_l^2} = -\left(\frac{\mathcal{L}_{li}}{\mathcal{L}_i} - \frac{\mathcal{L}_{hi}}{\mathcal{L}_i}\right)^2 \tag{8.4.21}$$

$$\frac{\partial^2 \ell_i}{\partial v_l \partial p_l} = \frac{\mathcal{L}_{li}}{\mathcal{L}_i}\left(\frac{\delta_i}{v_l} - z\right)\left(1 - p_l \frac{\mathcal{L}_{li} - \mathcal{L}_{hi}}{\mathcal{L}_i}\right) \tag{8.4.22}$$

$$\begin{aligned}
\frac{\partial^2 \ell_i}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\lambda}'} &= \frac{p_l}{\mathcal{L}_i}\left[\frac{\partial \mathcal{L}_{li}}{\partial \lambda}(\delta_i \boldsymbol{d}_i - v_l \mu_i \boldsymbol{T}_i) - \mathcal{L}_{li} \mathrm{diag}\left(v_l \mu_i \mathrm{e}^{\boldsymbol{\lambda}} * \boldsymbol{T}_i\right)\right] + \\
&+ \frac{p_h}{\mathcal{L}_i}\left[\frac{\partial \mathcal{L}_{hi}}{\partial \lambda}(\delta_i \boldsymbol{d}_i - \mu_i \boldsymbol{T}_i) - \mathcal{L}_{hi} \mathrm{diag}\left(\mu_i \mathrm{e}^{\boldsymbol{\lambda}} * \boldsymbol{T}_i\right)\right] - \\
&- \frac{1}{\mathcal{L}_i^2}\left(\frac{\partial \mathcal{L}_i}{\partial \boldsymbol{\lambda}}\right)^2
\end{aligned} \tag{8.4.23}$$

$$\begin{aligned}
\frac{\partial^2 \ell_i}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}'} &= p_l \frac{\mathcal{L}_{li}}{\mathcal{L}_i}\left[(\delta_i - z_{li})^2 - z_{li}\right] \boldsymbol{x}_i \boldsymbol{x}_i' + p_h \frac{\mathcal{L}_{hi}}{\mathcal{L}_i}\left[(\delta_i - z_{hi})^2 - z_{hi}\right] \boldsymbol{x}_i \boldsymbol{x}_i' - \\
&- \frac{\partial \mathcal{L}_i}{\partial \boldsymbol{\gamma}} \frac{\partial \mathcal{L}_i}{\partial \boldsymbol{\gamma}'}
\end{aligned} \tag{8.4.24}$$

$$\begin{aligned}
\frac{\partial^2 \ell_i}{\partial \boldsymbol{\lambda} \boldsymbol{\gamma}'} &= \left[(\delta_i - z_{li})(\delta_i \boldsymbol{d}_i - z_{li}) - z_{li}\right] \frac{\mathcal{L}_{li}}{\mathcal{L}_i} \boldsymbol{x}_i' p_l + \\
&+ \left[(\delta_i - z_{hi})(\delta_i \boldsymbol{d}_i - z_{hi}) - z_{hi}\right] \frac{\mathcal{L}_{hi}}{\mathcal{L}_i} \boldsymbol{x}_i' p_h - \frac{1}{\mathcal{L}_i} \frac{\partial \mathcal{L}_i}{\partial \boldsymbol{\lambda}} \frac{\partial \mathcal{L}_i}{\partial \boldsymbol{\gamma}'}
\end{aligned} \tag{8.4.25}$$

$$\begin{aligned}
\frac{\partial^2 \ell_i}{\partial \boldsymbol{\lambda} \partial v_l} &= p_l \frac{\mathcal{L}_{li}}{\mathcal{L}_i}\left(\frac{\delta_i}{v_l} - z_{hi}\right)\left(\delta_i \boldsymbol{d}_i - v_l \mu_i \mathrm{e}^{\boldsymbol{\lambda}} * \boldsymbol{T}_i - \frac{1}{\mathcal{L}_i} \frac{\partial \mathcal{L}_i}{\partial \boldsymbol{\lambda}}\right) - \\
&- p_l \frac{\mathcal{L}_{li}}{\mathcal{L}_i} \mu_i \mathrm{e}^{\boldsymbol{\lambda}} * \boldsymbol{T}_i
\end{aligned} \tag{8.4.26}$$

$$\begin{aligned}
\frac{\partial^2 \ell_i}{\partial \boldsymbol{\lambda} \partial p_l} &= \frac{\mathcal{L}_{li}}{\mathcal{L}_i}\left(\delta_i d_i - v_l \mu_i \mathrm{e}^{\boldsymbol{\lambda}} * \boldsymbol{T}_i\right) - \frac{\mathcal{L}_{hi}}{\mathcal{L}_i}\left(\delta_i d_i - \mu_i \mathrm{e}^{\boldsymbol{\lambda}} * \boldsymbol{T}_i\right) - \\
&- \frac{\mathcal{L}_{li} - \mathcal{L}_{hi}}{\mathcal{L}_i} \frac{1}{\mathcal{L}_i} \frac{\partial \mathcal{L}_i}{\partial \boldsymbol{\lambda}}
\end{aligned} \tag{8.4.27}$$

$$\begin{aligned}
\frac{\partial^2 \ell_i}{\partial \boldsymbol{\gamma} \partial v_l} &= p_l \frac{\mathcal{L}_{li}}{\mathcal{L}_i}\left(\frac{\delta_i}{v_l} - z_{hi}\right)(\delta_i - z_{li}) \boldsymbol{x}_i - \\
&- p_l \frac{\mathcal{L}_{li}}{\mathcal{L}_i}\left(\frac{\delta_i}{v_l} - z_{hi}\right) \frac{1}{\mathcal{L}_i} \frac{\partial \mathcal{L}_i}{\partial \boldsymbol{\gamma}} \boldsymbol{x}_i - p_l \frac{\mathcal{L}_{li}}{\mathcal{L}_i} z_i \boldsymbol{x}_i
\end{aligned} \tag{8.4.28}$$

$$\frac{\partial^2 \ell_i}{\partial \boldsymbol{\gamma} \partial p_l} = \left[\frac{\mathcal{L}_{li}}{\mathcal{L}_i}(\delta_i - z_{li}) - \frac{\mathcal{L}_{hi}}{\mathcal{L}_i}(\delta_i - z_{hi})\right] \boldsymbol{x}_i - \frac{\mathcal{L}_{li} - \mathcal{L}_{hi}}{\mathcal{L}_i} \frac{1}{\mathcal{L}_i} \frac{\partial \mathcal{L}_i}{\partial \boldsymbol{\gamma}} \tag{8.4.29}$$

Eelnevas tähendab $\{\mathrm{e}^{\boldsymbol{\lambda}}\}_i = \mathrm{e}^{\lambda_i}$, $*$ vektorite elementide kaupa korrutamist ($\{\boldsymbol{a} * \boldsymbol{b}\}_i = a_i b_i$) ning $\mathrm{diag}\, \boldsymbol{a}$ on maatriks, mille peadiagonaalil on vektor $\boldsymbol{a}$ ja mujal nullid.

Indiviidi-spetsiifiline heterogeensus Eeldame nii nagu eespool et hasart on konstantne iga M ajavahemiku sees, erinevatel ajavahemikel võib ta aga olla erinev. Avaldugu hasart

$$\theta(t|\boldsymbol{x}, v) = v e^{\lambda(t)} e^{\gamma x_i'}. \tag{8.4.30}$$

Mittevaadeldav heterogeensus olgu diskreetse jaotusega $v \in \{v_1, v_2, \ldots, v_K\}$ ja tõenäosusega vastavalt $p_1, p_2, \ldots, p_K$. Olgu inimese mittevaadeldav tunnus v ajas muutumatu, vaadeldav tunnus aga võib muutuda. Inimese i spelli j osa laiklihuudi funktsioonis on siis:

$$\mathcal{L}_{ij}(\cdot|v) = v\theta(t_{ij}|\boldsymbol{x}_{ij})^{\delta_{ij}} S(t_{ij}|\boldsymbol{x}_{ij}, v). \tag{8.4.31}$$

Siin t_{ij} on spelli vaadeldud kestus, δ_{ij} on mitte-tsenseerituse indikaator ja $\boldsymbol{x}_{ij}$ on inimese i vaadeldavad isikutunnused spelli j ajal. $\mathcal{L}_{ij}(\cdot|v)$ on analoogne indiviidi-spetsiifilise laiklihuudiga sõltumatute vaatluste juhul.

Olgu inimese i kohta N_i vaadeldud spelli. Inimese i osa laiklihuudis avaldub siis

$$\mathcal{L}_i(\cdot|v) = \prod_{j=1}^{N_i} \mathcal{L}_{ij}(\cdot|v). \tag{8.4.32}$$

Vaadeldav laiklihuud avaldub:

$$\mathcal{L}_i(\cdot) = \sum_{k=1}^{K} p_k \mathcal{L}_i(\cdot|v_k). \tag{8.4.33}$$

Log-laiklihuudi gradient avaldub:

$$\frac{\partial \ell_i}{\partial v_k} = \frac{p_k}{\mathcal{L}_i(\cdot)} \sum_j \frac{\mathcal{L}_i(\cdot|v_k)}{\mathcal{L}_{ij}(\cdot|v_k)} \frac{\partial \mathcal{L}_{ij}(\cdot|v_k)}{\partial v_k} \tag{8.4.34}$$

$$\frac{\partial \ell_i}{\partial p_l} = \frac{\mathcal{L}_i(\cdot|v_k)}{\mathcal{L}_i(\cdot)} \tag{8.4.35}$$

$$\frac{\partial \ell_i}{\partial \boldsymbol{\lambda}} = \frac{1}{\mathcal{L}_i(\cdot)} \sum_k p_k \sum_j \frac{\mathcal{L}_i(\cdot|v_k)}{\mathcal{L}_{ij}(\cdot|v_k)} \frac{\partial \mathcal{L}_{ij}(\cdot|v_k)}{\partial \boldsymbol{\lambda}} \tag{8.4.36}$$

$$\frac{\partial \ell_i}{\partial \boldsymbol{\gamma}} = \frac{1}{\mathcal{L}_i(\cdot)} \sum_k p_k \sum_j \frac{\mathcal{L}_i(\cdot|v_k)}{\mathcal{L}_{ij}(\cdot|v_k)} \frac{\partial \mathcal{L}_{ij}(\cdot|v_k)}{\partial \boldsymbol{\gamma}}. \tag{8.4.37}$$

$\mathcal{L}_{ij}(\cdot|v)$ tuletised on analoogilised nagu sõltumatute vaatluste korral. Lisaks tuleb p järgi gradiendi võtmisel arvestada, et $\sum p_k = 1$.

8.4.4 Intervallandmed

Intervallandmetega on tegemist siis kui on vaadeldav ainult fakt et sündmus (seisundite vahetamine, tsenseerimine) toimus mingis kestuse(aja)vahemikus (näiteks kuu või nädala jooksul). Intervallmudel sobib ka siis kui ei soovi hasarti täpselt spetsifitseerida.

Mudel: olgu kestus jagatud $T+1$ vahemikuks: $[0, t_1), [t_1, t_2), \ldots, [t_{T-1}, t_T), [t_T, \infty)$. Iga indiviidi i kohta olgu vaadeldav et millises vahemikus ta algsest seisundist lahkus, või et millises vahemikust vaatlus on tsenseeritud. Eeldame MPH mudelit nagu 8.4.3. osas:

$$\vartheta(\tau|\boldsymbol{x}, v) = \lambda(\tau)e^{\boldsymbol{\beta}'\boldsymbol{x}}v. \tag{8.4.38}$$

Tõenäosus, et isik jääb kogu intervalli n jooksul algseisundisse avaldub

$$S_n(\boldsymbol{x}, v) = \exp(-vz_n(\boldsymbol{x})) = \exp(-ve^{\boldsymbol{\beta}'\boldsymbol{x}} \int_{\tau_{n-1}}^{\tau_n} \lambda(s)\,\mathrm{d}s). \tag{8.4.39}$$

Nüüd võib defineerida $\tilde{\lambda}_n$:

$$e^{\tilde{\lambda}_n}(t_n - t_{n-1}) \equiv \int_{\tau_{n-1}}^{\tau_n} \lambda(s)\,\mathrm{d}s, \tag{8.4.40}$$

kus $e^{\tilde{\lambda}_s}$ on keskmine põhihasart vahemikus s ja $\tilde{\lambda}$ on lihtsalt mudeli parameeter. Seega põhihasart on spetsifitseeritud mitteparameetriliselt. Vaatluse laiklihuud fikseeritud v korral avaldub nüüd

$$\mathcal{L}(n|\boldsymbol{x}, v) = \left(1 - e^{-vz_n(\boldsymbol{x})}\right)^{\delta} \prod_{m=1}^{n-1} e^{-vz_m(\boldsymbol{x})}, \tag{8.4.41}$$

kus $\delta = 1$ tähendab et vaatlus pole tsenseeritud. Kogu vaatluse log-laiklihuud on

$$\ell(n|\boldsymbol{x}) = \log\left(\sum_{k=1}^{K} p_k \mathcal{L}(n|\boldsymbol{x}, v_k)\right). \tag{8.4.42}$$

v-spetsiifilise laiklihuudi gradient avaldub

$$\frac{\partial}{\partial\boldsymbol{\beta}}\mathcal{L}(n|\boldsymbol{x}, v) = vE_g(n|\boldsymbol{x}, v)\frac{\partial}{\partial\boldsymbol{\beta}}z_n(\boldsymbol{x}) - vS_g(n|\boldsymbol{x}, v)\sum_{m=1}^{n-1}\frac{\partial}{\partial\boldsymbol{\beta}}z_m(\boldsymbol{x}) \tag{8.4.43}$$

$$\frac{\partial}{\partial\tilde{\lambda}_s}\mathcal{L}(n|\boldsymbol{x}, v) = vE_g(n|\boldsymbol{x}, v)\frac{\partial}{\partial\tilde{\lambda}_s}z_n(\boldsymbol{x}) - vS_g(n|\boldsymbol{x}, v)\sum_{m=1}^{n-1}\frac{\partial}{\partial\tilde{\lambda}_s}z_m(\boldsymbol{x}) \tag{8.4.44}$$

$$\frac{\partial}{\partial v}\mathcal{L}(n|\boldsymbol{x}, v) = E_g(n|\boldsymbol{x}, v)z_n(\boldsymbol{x}) - S_g(n|\boldsymbol{x}, v)\sum_{m=1}^{n-1} z_m(\boldsymbol{x}) \tag{8.4.45}$$

kus

$$E_g(n|\boldsymbol{x}, v) = \delta e^{-vz_n(\boldsymbol{x})} \prod_{m=1}^{n-1} e^{-vz_m(\boldsymbol{x})} \tag{8.4.46}$$

$$S_g(n|\boldsymbol{x}, v) = \mathcal{L}(n|\boldsymbol{x}, v) = \left(1 - \mathrm{e}^{-vz_n(\boldsymbol{x})}\right)^{\delta} \prod_{m=1}^{n-1} \mathrm{e}^{-vz_m(\boldsymbol{x})} \tag{8.4.47}$$

on gradiendi seisundist lahkumise ja seisundis kestmise spetsiifilised osad ja z gradient avaldub

$$\frac{\partial}{\partial \boldsymbol{\beta}} z_n(\boldsymbol{x}) = z_n(\boldsymbol{x}) \boldsymbol{x}_n \tag{8.4.48}$$

$$\frac{\partial}{\partial \tilde{\lambda}_s} z_n(\boldsymbol{x}) = z_n(\boldsymbol{x}) \mathbb{1}(s = n). \tag{8.4.49}$$

Kogulaiklihuudi gradient on

$$\frac{\partial}{\partial \boldsymbol{\beta}} \ell(n|\boldsymbol{x}) = \frac{1}{\mathcal{L}(n|\boldsymbol{x})} \left(\sum_{k=1}^{K} p_k \frac{\partial}{\partial \boldsymbol{\beta}} \mathcal{L}(n|\boldsymbol{x}, v) \right) \tag{8.4.50}$$

$$\frac{\partial}{\partial \tilde{\lambda}_s} \ell(n|\boldsymbol{x}) = \frac{1}{\mathcal{L}(n|\boldsymbol{x})} \left(\sum_{k=1}^{K} p_k \frac{\partial}{\partial \tilde{\lambda}_s} \mathcal{L}(n|\boldsymbol{x}, v) \right) \tag{8.4.51}$$

$$\frac{\partial}{\partial v_k} \ell(n|\boldsymbol{x}) = \frac{1}{\mathcal{L}(n|\boldsymbol{x})} p_k \frac{\partial}{\partial v_k} \mathcal{L}(n|\boldsymbol{x}, v) \tag{8.4.52}$$

$$\frac{\partial}{\partial p_k} \ell(n|\boldsymbol{x}) = \frac{\mathcal{L}(n|\boldsymbol{x}, v_k)}{\mathcal{L}(n|\boldsymbol{x})} \tag{8.4.53}$$

Mitu lѳppseisundit Mitme lѳppseisundiga mudelid esitatakse sageli *kompiiting risk* kujul. Too tähendab et kujutatakse ette et kѳikidesse lѳppseisunditesse viivad sѳltumatud Markovi protsessid, realiseerub too mille aeg on kѳige lühem. Vajadusel vѳib tsenseerimist kujutada ühena lѳppseisunditest.

Kui kѳik muutujad on vaadeldavad, ei erine mitme lѳppseisundiga juhtub olukorrast kui modelleerida üksikuid lѳppseisundeid sѳltumatult, teised seisundid oleksid siis nagu tsenseeritud. Kui mudelis on mittevaadeldav heterogeensus, on pilt ainult veidi keerulisem.

Olgu M vѳimalikku lѳppseisundit ja $m \in \{1, \ldots, M\}$ olgu lѳppseisundi näitaja. Olgu seisundisse m ülemineku hasart, sѳltuvalt vaadeldavatest ja mittevaadeldavatest parameetritest, $\vartheta^m(\tau|\boldsymbol{x}, v^m)$. Mittevaadeldav heterogeensus olgu M-mѳѳtmelise diskreetse jaotusega: $v^m \in \{v_1^m, \ldots, v_{K^m}^m$ kusjuures $\boldsymbol{v} = (v_{k^1}^1, \ldots, v_{k^M}^M)$ esineb tѳenäosusega $p_{k^1 \ldots k^M}$. Eeldame et erinevate spellide jooksul on $\boldsymbol{v}$ konstantne, $\boldsymbol{x}$ aga vѳib muutuda.

Inimese i spelli j, osa laiklihuudi funktsioonis siirde m järgi on nüüd:

$$\mathcal{L}_{ij}^m(\cdot|v^m) = \left[\vartheta^m(\tau_{ij}|\boldsymbol{x}_{ij})\right]^{\delta_{ij}^m} S^m(\tau_{ij}|\boldsymbol{x}_{ij}, v^m). \tag{8.4.54}$$

Siirdega seotud liige $\left[\vartheta^m(\tau_{ij}|\boldsymbol{x}_{ij})\right]^{\delta_{ij}^m}$ ja püsimisega seotud liige $S^m(\tau_{ij}|\boldsymbol{x}_{ij}, v^m)$ avalduvad nii nagu ühe spelli ja ühe lѳppseisundi puhul (vaata näiteks 8.4.3 vѳi 8.4.4). Siin τ_{ij} on spelli vaadeldud kestus, δ_{ij}^m on mitte-tsenseerituse indikaator m-seisundi mѳttes (= 1 kui läks seisundisse m ja 0 kui es lähe) ja $\boldsymbol{x}_{ij}$ on inimese i vaadeldavad isikutunnused spelli j ajal. Spelli kogulaiklihuud on:

$$\mathcal{L}_{ij}(\cdot|\boldsymbol{x}_{ij}, \boldsymbol{v}) = \prod_{m=1}^{M} \mathcal{L}_{ij}^m(\cdot|\boldsymbol{x}_{ij}, v^m). \tag{8.4.55}$$

Kompiiting risk mudel eeldab et siirded on sѳltumatud (kui kontrollida $\boldsymbol{x}$ ja $\boldsymbol{v}$ suhtes), seega siis laiklihuudi korrutis siirete kaupa.

Olgu inimese i kohta N_i vaadeldud spelli. Inimese i osa laiklihuudis avaldub siis

$$\mathcal{L}_i(\cdot|\boldsymbol{x}, \boldsymbol{v}) = \prod_{j=1}^{N_i} \mathcal{L}_{ij}(\cdot|\boldsymbol{x}, \boldsymbol{v}). \tag{8.4.56}$$

ja vaadeldav laiklihuud avaldub:

$$\mathcal{L}_i(\cdot|\boldsymbol{x}) = \sum_{k^1=1}^{K^1} \cdots \sum_{k^M=1}^{K^M} p_{k^1 \ldots k^M} \mathcal{L}_i(\cdot|\boldsymbol{x}, \boldsymbol{v}). \tag{8.4.57}$$

Laiklihuudi gradient avaldub:

$$\frac{\partial \mathcal{L}_i(\cdot|\boldsymbol{x})}{\partial \boldsymbol{\lambda}^m} = \sum_{k^1=1}^{K^1} \cdots \sum_{k^M=1}^{K^M} p_{k^1 \ldots k^M} \mathcal{L}_i(\cdot|\boldsymbol{x}, \boldsymbol{v}) \sum_{j=1}^{N_i} \left[\frac{1}{\mathcal{L}_{ij}^m(\cdot|\boldsymbol{x} v_{k^m}^m)} \frac{\partial \mathcal{L}_{ij}^m(\cdot|\boldsymbol{x} v_{k^m}^m)}{\partial \boldsymbol{\lambda}^m} \right] \tag{8.4.58}$$

$$\frac{\partial \mathcal{L}_i(\cdot|\boldsymbol{x})}{\partial \boldsymbol{\gamma}^m} = \sum_{k^1=1}^{K^1} \cdots \sum_{k^M=1}^{K^M} p_{k^1 \ldots k^M} \mathcal{L}_i(\cdot|\boldsymbol{x}, \boldsymbol{v}) \sum_{j=1}^{N_i} \left[\frac{1}{\mathcal{L}_{ij}^m(\cdot|\boldsymbol{x} v_{k^m}^m)} \frac{\partial \mathcal{L}_{ij}^m(\cdot|\boldsymbol{x} v_{k^m}^m)}{\partial \boldsymbol{\gamma}^m} \right] \tag{8.4.59}$$

$$\frac{\partial \mathcal{L}_i(\cdot|\boldsymbol{x})}{\partial v_k^m} = \sum_{k^1=1}^{K^1} \cdots \sum_{k^{m-1}=1}^{K^{m-1}} \sum_{k^{m+1}=1}^{K^{m+1}} \cdots \sum_{k^M=1}^{K^M} p_{k^1 \ldots k^{m-1} k k^{m+1} \ldots k^M} \cdot$$

$$\cdot \mathcal{L}_i(\cdot|\boldsymbol{x}, \boldsymbol{v}) \sum_{j=1}^{N_i} \left[\frac{1}{\mathcal{L}_{ij}^m(\cdot|\boldsymbol{x} v_k^m)} \frac{\partial \mathcal{L}_{ij}^m(\cdot|\boldsymbol{x} v_k^m)}{\partial \gamma^m} \right] \tag{8.4.60}$$

$$\frac{\partial \mathcal{L}_i(\cdot|\boldsymbol{x})}{\partial p_{k^1 \dots k^M}} = \mathcal{L}_i(\cdot|\boldsymbol{x}, (v_{k^1}^1, \dots, v_{k^M}^M)) \tag{8.4.61}$$

$\mathcal{L}_{ij}^m(\cdot|\boldsymbol{x}, \boldsymbol{v})$ tuletised on nii nagu sõltumatute vaatluste korral.

8.4.5 Parametriseerimine

Logistic transition probability in discrete-time In discrete time, it is useful to parameterise the destination-specific exit probabilities logistically as

$$p^m \equiv \Pr(\text{exit to destination } m) = \frac{e^{\lambda^m}}{1 + \sum_k e^{\lambda^k}}. \tag{8.4.62}$$

The probabilities are guaranteed to be positive and their sum to be less than one while λ-s may be unbounded. The components of the corresponding gradient transformation matrix are

$$\frac{\partial}{\partial \lambda^k} p^m = \mathbb{1}(k = m) p_m - p_k p_m \tag{8.4.63}$$

and the inverse transformation

$$l^m = \log \frac{p^m}{1 - \sum_k p^k}. \tag{8.4.64}$$

Transition probability in continuous time A good choice for parametrising the time-dependent part of the hazard is

$$\lambda = e^{\tilde{\lambda}} \tag{8.4.65}$$

The λ is now guaranteed to be positive.

Diskreetne mittevaadeldav heterogeensus Diskreetne mittevaadeldav heterogeensus, üks lõppseisund: $v \in \{v_1, v_2, \dots, v_K\}$ ja vastavad tõenäousused $p_1, p_2, \dots, p_K$.

Kui laiklihuudi maksimeerimisel kasutada vastavaid parameetreid $\tilde{v}$ ning $\tilde{p}$ ($\tilde{N}$ parameetrit) kuid laiklihuudi arvutamiseks nad normaalkujule (N parameetrit) teisendada, siis peab arvestama, et vastavad gradiendi komponendid teisenevad:

$$\begin{pmatrix} \frac{\partial}{\partial \tilde{\boldsymbol{v}}} \ell_i \\ \frac{\partial}{\partial \tilde{\boldsymbol{p}}} \ell_i \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial \tilde{\boldsymbol{v}}} v & \frac{\partial}{\partial \tilde{\boldsymbol{v}}} p \\ \frac{\partial}{\partial \tilde{\boldsymbol{p}}} v & \frac{\partial}{\partial \tilde{\boldsymbol{p}}} p \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial v} \ell_i \\ \frac{\partial}{\partial p} \ell_i \end{pmatrix} \equiv \mathsf{C} \begin{pmatrix} \frac{\partial}{\partial v} \ell_i \\ \frac{\partial}{\partial p} \ell_i \end{pmatrix} \tag{8.4.66}$$

Maatriksi C ridade arv vastab algkuju parameetrite arvule $\tilde{N}$ ning veergude arv normaalkuju parameetrite arvule N (s.h. lineaarselt sõltuvad p_H ja v_H). Kovarjatsjoonimaatriksi p ning v sisaldav osa on vastavalt:

$$\mathsf{\Sigma} = \mathsf{C}' \tilde{\mathsf{\Sigma}} \mathsf{C} \tag{8.4.67}$$

Heterogeensus Tõenäosused on võimalik parametriseerida kui

$$p_k = \frac{e^{\tilde{p}_k}}{\sum_{k=1}^{K} e^{\tilde{p}_k}} \qquad \text{ja} \qquad \sum_{k=1}^{K} \tilde{p}_k = 0, \tag{8.4.68}$$

kus K on jaotuse toetuspunktide arv.

Keskväärtuse normeerimine v keskväärtuse võib normeerida üheks:

$$v_k = \mathrm{e}^{\tilde{v}_k} \qquad \text{ja} \qquad \sum_{k=1}^{K} p_k v_k = 1. \tag{8.4.69}$$

Vastav pöördteisendus tõenäosuste tarvis on

$$\tilde{p}_k = \log p_k - \frac{\sum_{l=1}^{K} \log p_l}{K}. \tag{8.4.70}$$

C komponendid on:

$$\frac{\partial}{\partial \tilde{v}_l} v_k = \begin{cases} v_k & \text{kui} \qquad k = l < K \\ -v_l \dfrac{p_l}{p_K} & \text{kui} \qquad k = K \\ 0 & \text{muudel juhtudel} \end{cases} \tag{8.4.71}$$

$$\frac{\partial}{\partial \tilde{v}_l} p_k = 0 \tag{8.4.72}$$

$$\frac{\partial}{\partial \tilde{p}_l} v_k = \begin{cases} \dfrac{1}{p_K} \left[(1 - p_K) + p_K v_K + (1 - p_l) v_l \right] & \text{kui} \qquad k = K \\ 0 & \text{muudel juhtudel} \end{cases} \tag{8.4.73}$$

$$\frac{\partial p_k}{\partial \tilde{p}_l} = -p_k (p_l - p_K) + \mathbb{1}(k = l) p_K - \mathbb{1}(K = k) p_K \tag{8.4.74}$$

v **komponendi normeerimine** Defineerime

$$v_K = 1. \tag{8.4.75}$$

Näib et nii on intervallandmete juures tulemused mõnevõrra stabiilsemad, samas on põhihasarti raskem tõlgendada.

C komponendid on

$$\frac{\partial}{\partial \tilde{v}_l} v_k = \begin{cases} v_k & \text{kui} \qquad k = l < K \\ 0 & \text{muudel juhtudel} \end{cases} \tag{8.4.76}$$

$$\frac{\partial}{\partial \tilde{v}_l} p_k = 0 \tag{8.4.77}$$

$$\frac{\partial}{\partial \tilde{p}_l} v_k = 0 \tag{8.4.78}$$

$$\frac{\partial p_k}{\partial \tilde{p}_l} = -\frac{\mathrm{e}^{\tilde{p}_k} \left(\mathrm{e}^{\tilde{p}_l} - \mathrm{e}^{\tilde{p}_H} \right)}{\left(\sum_{i=1}^{H} \mathrm{e}^{\tilde{p}_i} \right)^2} + \mathbb{1}(k = l) \frac{\mathrm{e}^{\tilde{p}_k}}{\sum_{i=1}^{H} \mathrm{e}^{\tilde{p}_i}} + \mathbb{1}(K = k) \frac{\mathrm{e}^{\tilde{p}_K}}{\sum_{i=1}^{K} \mathrm{e}^{\tilde{p}_i}} \tag{8.4.79}$$

M **lõppseisundit** (8.4.66) asemel võib nüüd kirjutada:

$$\begin{pmatrix} \frac{\partial}{\partial \tilde{\boldsymbol{v}}^1}\ell_i \\ \frac{\partial}{\partial \tilde{\boldsymbol{v}}^2}\ell_i \\ \cdots \\ \frac{\partial}{\partial \tilde{\boldsymbol{v}}^M}\ell_i \\ \frac{\partial}{\partial \tilde{\boldsymbol{p}}}\ell_i \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial \tilde{\boldsymbol{v}}^1}v^1 & \frac{\partial}{\partial \tilde{\boldsymbol{v}}^1}v^2 & \cdots & \frac{\partial}{\partial \tilde{\boldsymbol{v}}^1}v^M & \frac{\partial}{\partial \tilde{\boldsymbol{v}}^1}p \\ \frac{\partial}{\partial \tilde{\boldsymbol{v}}^2}v^1 & \frac{\partial}{\partial \tilde{\boldsymbol{v}}^2}v^2 & \cdots & \frac{\partial}{\partial \tilde{\boldsymbol{v}}^2}v^M & \frac{\partial}{\partial \tilde{\boldsymbol{v}}^2}p \\ \cdots & \cdots & \ddots & \cdots & \cdots \\ \frac{\partial}{\partial \tilde{\boldsymbol{v}}^M}v^1 & \frac{\partial}{\partial \tilde{\boldsymbol{v}}^M}v^2 & \cdots & \frac{\partial}{\partial \tilde{\boldsymbol{v}}^M}v^M & \frac{\partial}{\partial \tilde{\boldsymbol{v}}^M}p \\ \frac{\partial}{\partial \tilde{\boldsymbol{p}}}v^1 & \frac{\partial}{\partial \tilde{\boldsymbol{p}}}v^2 & \cdots & \frac{\partial}{\partial \tilde{\boldsymbol{p}}}v^M & \frac{\partial}{\partial \tilde{\boldsymbol{p}}}p \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial \boldsymbol{v}^1}\ell_i \\ \frac{\partial}{\partial \boldsymbol{v}^2}\ell_i \\ \cdots \\ \frac{\partial}{\partial \boldsymbol{v}^M}\ell_i \\ \frac{\partial}{\partial \boldsymbol{p}}\ell_i \end{pmatrix}$$

$$\equiv \mathsf{C} \begin{pmatrix} \frac{\partial}{\partial \boldsymbol{v}^1}\ell_i \\ \frac{\partial}{\partial \boldsymbol{v}^2}\ell_i \\ \cdots \\ \frac{\partial}{\partial \boldsymbol{v}^M}\ell_i \\ \frac{\partial}{\partial \boldsymbol{p}}\ell_i \end{pmatrix}, \quad (8.4.80)$$

kus C on $\tilde{N} \times N$ maatriks.

Olgu p^m_{Dk} suuna m spetsiifiline tøenäosus, s.t. millise tøenäosusega esineb väärtus v^m_k. Tuletised vøib nüüd avaldada kui

$$\frac{\partial}{\partial \tilde{p}}v^m = \frac{\partial}{\partial \tilde{p}}p \frac{\partial}{\partial p}p^m_D \frac{\partial}{\partial p^m_D}v^m \quad (8.4.81)$$

9 Algorithms

9.1 Arrays

Array indexing Let M be a N-dimensional column-major zero-based array, and $m[i_1, i_2, \ldots, i_N]$ its element where $i_j \in \{0, \ldots, K_j - 1\}$ and K_j is the size of dimension j. Vectorized index for element $m[i_1, i_2, \ldots, i_N]$ is

$$j = i_1 + K_1 i_2 + K_2 i_3 + \cdots + K_{N-1} i_N. \tag{9.1.1}$$

The array indices can be recovered from the vector index j as:

$$\begin{aligned}
i_1 &= j \mod K_1 \\
i_2 &= [j \mod (K_1 K_2)] \,//\, K_1 \\
i_3 &= [j \mod (K_1 K_2 K_3)] \,//\, (K_1 K_2) \\
i_4 &= [j \mod (K_1 K_2 K_3 K_4)] \,//\, (K_1 K_2 K_3) \\
&\cdots
\end{aligned}$$

where $//$ is the integer division.

References

Calzolari, G. and Fiorentini, G. (1993) Alternative covariance estimators of the standard Tobit model, *Economics Letters*, **42**, 5–13.

Chiang, A. C. (1984) *Fundamental Methods of Mathematical Economics*, McGraw-Hill, Singapore, third edn.

Lin, J. (1991) Divergence measures based on the shannon entropy, *Information Theory, IEEE Transactions on*, **37**, 145–151.

Miller, R. E. (1979) *Dynamic Optimization and Economic Applications*, McGraw-Hill, New York.