5. Vorlesung, 27. Oktober 2014 nach D.S. Sivia, Data Analysis, Oxford Univ. Press, Chapter 3.6

Thema: $\boxed{\text{Fehlerfortpflanzung}}$

Fragestellung: Wir haben eine Posteriorverteilung $\mathrm{pdf}(X,Y\,|\,I)$ für die gesuchten Parameter X, Y berechnet.
Wie finden wir $\mathrm{pdf}(Z\,|\,I)$, wobei $Z = h(X,Y)$?

Beispiel: Bei der Messung des Halleffekts haben wir die Steigung R_H der Hallgeraden bestimmt (lin. Regression)
Wie gross ist die Elektronendichte $n_s = \frac{1}{R_H e}$? Wie gross ist die Standardabweichung für n_s?

1. Fall: Nur ein Parameter

Gegeben: $\mathrm{pdf}(X\,|\,I) = f(X)$, gesucht $\mathrm{pdf}(Z\,|\,I) = g(z)$, wobei $Z = h(X)$

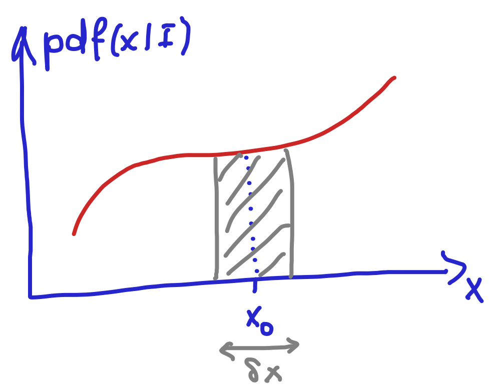
Jedem Wert x_0 entspricht ein korrespondierender Wert z_0. Betrachten wir ein kleines Intervall $[x_0 - \frac{\delta x}{2}, x_0 + \frac{\delta x}{2}]$, dann ist die Wahrscheinlichkeit, dass x in diesem Intervall liegt, gegeben durch

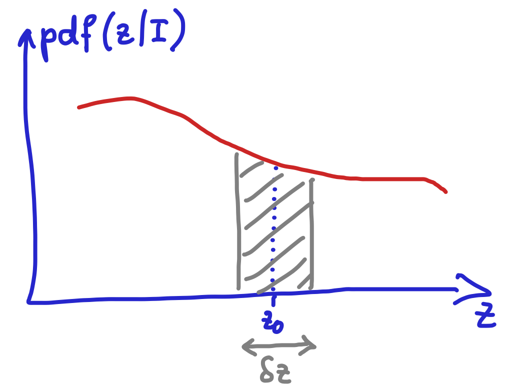
$$\mathrm{prob}\left(x_0 - \frac{\delta x}{2} \le x \le x_0 + \frac{\delta x}{2}\,\Big|\, I\right) = \int_{x_0-\delta x/2}^{x_0+\delta x/2} dx\, \mathrm{pdf}(x\,|\,I) \approx \mathrm{pdf}(x_0\,|\,I)\,\delta x$$

Entsprechend ist

$$\mathrm{prob}\left(z_0 - \frac{\delta z}{2} \le z \le z_0 + \frac{\delta z}{2}\,\Big|\, I\right) \approx \mathrm{pdf}(z_0\,|\,I)\,\delta z$$

Graphisch:





Wir fordern: $\mathrm{pdf}(x\,|\,I)\,\delta x = \mathrm{pdf}(z\,|\,I)\,\delta z \Rightarrow f(x) = g[h(x)] \times \left|\frac{dh(x)}{dx}\right|$

Das ist ein einfaches Beispiel einer Koordinatentransformation.

Beispiel: Gegeben sei die Gammaverteilung

$$\Gamma(x;\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-x/\beta} \quad \text{für } x>0$$

Wie lautet die Verteilung für $z = \frac{1}{x}$ für $z>0$?

$$f(x)dx = \frac{1}{\Gamma(\alpha)\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-x/\beta}\,dx = g(z)\,dz$$

$$g(z) = \frac{1}{\Gamma(\alpha)\beta}\left(\frac{1}{\beta z}\right)^{\alpha-1} e^{-1/\beta z} \times \left|-\frac{1}{z^2}\right| = \frac{1}{\Gamma(\alpha)\tilde\beta}\left(\frac{\tilde\beta}{z}\right)^{\alpha+1} e^{-\tilde\beta/z} \quad \text{mit } \tilde\beta := 1/\beta$$

Das ist die inverse Gammaverteilung $\mathrm{inv}\Gamma(x;\alpha,\tilde\beta)$.

2. Fall: M Parameter werden in M andere Parameter transformiert

$$\mathrm{pdf}(x_1,x_2,\ldots x_M)\,\delta x_1 \delta x_2 \ldots \delta x_M = \mathrm{pdf}(z_1,z_2,\ldots,z_M)\,\delta^M \mathrm{Vol}(\{z_j\})$$

Hier ist $\delta^M \mathrm{Vol}(\{z_j\})$ das infinitesimale Volumenelement, das die Abbildung des Hyperkubus $\delta x_1 \ldots \delta x_M$ ist.

Man findet: $$\mathrm{pdf}(x_1\ldots x_M) = \mathrm{pdf}(z_1\ldots z_M) \times \underbrace{\det\left(\frac{\partial(z_1\ldots z_M)}{\partial(x_1\ldots x_M)}\right)}_{\text{Determinante der Jacobi-Matrix}}$$

Beispiel: Gegeben sei die zweidim. Normalverteilung

$$pdf(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$

Wie lautet die Verteilung in Polarkoordinaten (R,θ)?

$x = R\cos\theta \;,\; y = R\sin\theta$

$$\det\left(\frac{\partial(x,y)}{\partial(R,\theta)}\right) = \begin{vmatrix} \cos\theta & -R\sin\theta \\ \sin\theta & R\cos\theta \end{vmatrix} = R(\cos^2\theta + \sin^2\theta) = R$$

$$\Rightarrow pdf(R,\theta) = pdf(x,y) \times R = \frac{R}{2\pi\sigma^2} \exp\left(-\frac{R^2}{2\sigma^2}\right)$$

3. Fall: wir transformieren von M auf N<M Parameter:

→ transformiere zunächst von M auf M Parameter (siehe oben)

→ marginalisiere die überschüssigen Parameter

Beispiel: wie oben, $pdf(x,y)$ ist zweidim. Normalverteilung für die kartes. Koord. x, y

Wie lautet die Verteilung für R?

→ transformiere von $(x,y) \to (R,\theta)$ (siehe oben)

→ marginalisiere θ:

$$pdf(R) = \int_0^{2\pi} pdf(R,\theta)\, d\theta = \frac{R}{\sigma^2} \exp\left(-\frac{R^2}{2\sigma^2}\right)$$

Mathematisch formale Prozedur:

Damit haben wir die zentralen Elemente der Fehlerfortpflanzung kennen gelernt: Koordinatentransformation und Marginalisierung, oder eine Kombination der beiden. Mathematisch formal lässt sich das folgendermassen ausdrücken:

$$\boxed{pdf(z|I) = \int dx \int dy\; pdf(x,y,z|I) = \iint dx\,dy\; pdf(x,y|I)\, pdf(z|x,y) = \iint dx\,dy\; pdf(x,y|I)\, \delta(z-h(x,y))} \quad (*)$$

Das lässt sich leicht auf Verteilungen höherer Dimensionen verallgemeinern.

Beispiel: $z = x+y$ und $pdf(x,y|I) = \underbrace{pdf(x|I)}_{f(x)}\underbrace{pdf(y|I)}_{g(y)}$ seien gegeben. Dann ist

$$pdf(z|I) = \int dx \int dy\; pdf(x,y|I)\, \delta(z-x-y) = \int dx \int dy\; f(x)g(y)\delta(z-x-y) = \int dx\, f(x)\, g(z-x)$$

Die Verteilung für die Summe zweier Parameter ist demnach die Faltung der beiden Verteilungen.

Nehmen wir im Speziellen an, das $f(x)$ und $g(y)$ Normalverteilungen

$f(x) = \mathcal{N}(x; x_0, \sigma_x)$ und $\mathcal{N}(y; y_0, \sigma_y)$ sind, dann ist

$$pdf(z=x+y|I) = \mathcal{N}\left(z; x_0+y_0, \sqrt{\sigma_x^2+\sigma_y^2}\right),$$

d.h. die Mittelwerte addieren sich: $z_0 = x_0 + y_0$ und die Fehler addieren sich quadratisch:

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2.$$

Eine nützliche Näherung

Betrachten wir Gl. (*) unter dem Aspekt, dass wir Mittelwert und Standardabweichung von z ausrechnen wollen.

Sei $pdf(x,y|I) = f(x,y)$ mit scharfem Maximum bei (x_0,y_0) und $z = h(x,y)$.

Das scharfe Maximum erlaubt uns eine Entwicklung von $\ln f(x,y)$ um (x_0,y_0) so dass in guter Näherung gilt:

$$\ln f(x,y) \approx \ln f(x_0,y_0) - \frac{1}{2}\begin{pmatrix}x\\y\end{pmatrix} H \begin{pmatrix}x\\y\end{pmatrix} \qquad \text{(siehe Näherung von Laplace)}$$

Die Matrix H ist die Hesse-Matrix der zweiten Ableitungen von $-\ln f(x,y)$ ausgewertet bei (x_0,y_0).
In dieser Näherung ist $f(x,y)$ gaussförmig (in 2 Dimensionen), so dass

$\langle x \rangle = x_0$ und $\langle y \rangle = y_0$

$\langle (x-x_0)^2 \rangle = \sigma_x^2 = \mathrm{Var}(x)$

$\langle (y-y_0)^2 \rangle = \sigma_y^2 = \mathrm{Var}(y)$

$\langle (x-x_0)(y-y_0) \rangle = \sigma_x \sigma_y \rho \equiv \mathrm{cov}(x,y)$ "Kovarianz von x und y"

Ebenso entwickeln wir

$$h(x,y) \approx h(x_0,y_0) + \left.\frac{\partial h}{\partial x}\right|_{x_0}(x-x_0) + \left.\frac{\partial h}{\partial y}\right|_{y_0}(y-y_0)$$

Damit ist

$$\langle z \rangle = \int dz\, z \int dx \int dy\, f(x,y)\, \delta(z-h(x,y))$$
$$= \int dx \int dy\, h(x,y) f(x,y) \approx \iint dx\,dy \left[h(x_0,y_0) + \left.\frac{\partial h}{\partial x}\right|_{x_0}(x-x_0) + \left.\frac{\partial h}{\partial y}\right|_{y_0}(y-y_0)\right] f(x,y)$$
$$\approx h(x_0,y_0)$$

$$\langle z^2 \rangle \approx \iint dx\,dy \left[h(x_0,y_0) + \left.\frac{\partial h}{\partial x}\right|_{x_0}(x-x_0) + \left.\frac{\partial h}{\partial y}\right|_{y_0}(y-y_0)\right]^2 f(x,y)$$
$$\approx h^2(x_0,y_0) + \left(\left.\frac{\partial h}{\partial x}\right|_{x_0}\right)^2 \langle (x-x_0)^2 \rangle + 2\left(\left.\frac{\partial h}{\partial x}\right|_{x_0}\right)\left(\left.\frac{\partial h}{\partial y}\right|_{y_0}\right)\langle (x-x_0)(y-y_0) \rangle + \left(\left.\frac{\partial h}{\partial y}\right|_{y_0}\right)^2 \langle (y-y_0)^2 \rangle$$

$$\langle (z-\langle z \rangle)^2 \rangle = \langle z^2 \rangle - \langle z \rangle^2 = \left(\left.\frac{\partial h}{\partial x}\right|_{x_0}\right)^2 \sigma_x^2 + 2\left(\left.\frac{\partial h}{\partial y}\right|_{y_0}\right)\left(\left.\frac{\partial h}{\partial x}\right|_{x_0}\right)\mathrm{cov}(x,y) + \left(\left.\frac{\partial h}{\partial y}\right|_{y_0}\right)^2 \sigma_y^2$$

$$\Rightarrow \boxed{\sigma_z = \sqrt{\left(\left.\frac{\partial h}{\partial x}\right|_{x_0}\right)^2 \sigma_x^2 + 2\left(\left.\frac{\partial h}{\partial x}\right|_{x_0}\right)\left(\left.\frac{\partial h}{\partial y}\right|_{y_0}\right)\mathrm{cov}(x,y) + \left(\left.\frac{\partial h}{\partial y}\right|_{y_0}\right)^2 \sigma_y^2}}$$

Gauss'sche Fehlerfortpflanzung

Beispiele:

1. $z = h(x,y) = x \pm y \Rightarrow \sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2}$ (siehe oben)

2. $z = h(x,y) = xy \Rightarrow \sigma_z = \sqrt{y_0^2 \sigma_x^2 + 2 x_0 y_0\, \mathrm{cov}(x,y) + x_0^2 \sigma_y^2}$

$$\Rightarrow \frac{\sigma_z}{z_0} = \sqrt{\left(\frac{\sigma_x}{x_0}\right)^2 + 2\,\frac{\mathrm{cov}(x,y)}{x_0 y_0} + \left(\frac{\sigma_y}{y_0}\right)^2}$$

3. $z = h(x,y) = \frac{x}{y} \Rightarrow \sigma_z = \sqrt{\frac{\sigma_x^2}{y_0^2} - 2\,\frac{x_0}{y_0^3}\,\mathrm{cov}(x,y) + \frac{x_0^2}{y_0^4}\,\sigma_y^2}$

$$\Rightarrow \frac{\sigma_z}{z_0} = \sqrt{\left(\frac{\sigma_x}{x_0}\right)^2 - 2\,\frac{\mathrm{cov}(x,y)}{x_0 y_0} + \left(\frac{\sigma_y}{y}\right)^2}$$