A primer in statistics – Random variables

Sensor fusion & nonlinear filtering

Lars Hammarstrand

DISCRETE-VALUED RANDOM VARIABLES

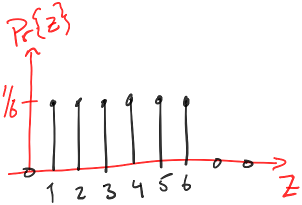
Probability mass function, pmf

- The *probability mass function* (pmf) of a discrete-valued random variable is denoted, $\Pr\{z\}$ or $P\{z\}$, where

$$\Pr\{z = i\} \geq 0 \qquad \text{for all } i$$
$$\sum_z \Pr\{z\} = 1.$$

Example: A fair dice

$$\Pr\{z = i\} = \begin{cases} \frac{1}{6} & \text{if } i = 1, 2, \ldots, 6 \\ 0 & \text{otherwise.} \end{cases}$$



CONTINUOUS-VALUED RANDOM VARIABLES

Probability density function (pdf)

- The *probability density function* (pdf) of a continuous-valued random variable is denoted $p(z)$, where

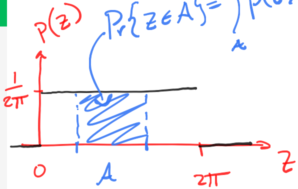
$$p(z) \geq 0 \text{ for all } z, \quad \text{and} \quad \int p(z)\, dz = 1$$

Example: Uniform distribution

- Suppose z is uniformly distributed between 0 and 2π, it's pdf is then

$$p(z) = \begin{cases} \frac{1}{2\pi} & \text{if } 0 \leq z < 2\pi \\ 0 & \text{otherwise.} \end{cases}$$

$\Pr\{z \in A\} = \int_A p(z)\, dz$



A primer in statistics – Conditional, Joint and marginal distributions

Sensor fusion & nonlinear filtering

Lars Hammarstrand

CONDITIONAL DISTRIBUTIONS

- Conditional distributions are indispensable components in sensor fusion, filtering and Bayesian estimation in general.

Conditional distribution (product rule)

- Let x and z be two random variables with the joint pdf $p(x, z)$.
- The *conditional density function*, $p(z|x)$, is defined through

$$p(x, z) = p(z|x)p(x),$$

and if $p(x) \neq 0$ this implies that

$$p(z|x) = \frac{p(x, z)}{p(x)}.$$

$p(z|x=x') = \frac{p(x',z)}{p(x')} \propto p(x',z)$

- Interpretation: $p(z|x)$ describes the distribution of z given that x is known.

CONDITIONAL DISTRIBUTIONS

Example: Candy problem

- Every day Sara decides how many pieces of candy she can have for an after lunch snack.
- With 40% probability she tosses a coin, heads means 1 piece and tails means 0 pieces
- With 60% probability she throws a dice (number on the dice = number of candies).

- If z denotes number of candies she eats

$$\Pr\left\{z = i \middle| \text{Sara tosses a coin}\right\} = \begin{cases} 0.5 & \text{if } i = 0, 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\Pr\left\{z = i \middle| \text{Sara throws a dice}\right\} = \begin{cases} 1/6 & \text{if } i = 1, 2, \dots, 6 \\ 0 & \text{otherwise} \end{cases}$$

LAW OF TOTAL PROBABILITY

- Many important results in non-linear filtering is obtained from the law of total probability.

Law of total probability (sum rule)

- If x takes values in a set S_x, the law of total probability states that

$$\text{Discrete:} \qquad \Pr\{z\} = \sum_{x \in S_x} \Pr\{x, z\} = \sum_{x \in S_x} \Pr\{z \mid x\} \Pr\{x\}$$

$$\text{Continuous:} \qquad p(z) = \int_{x \in S_x} p(x, z)\, dx = \int_{x \in S_x} p(z \mid x) p(x)\, dx$$

LAW OF TOTAL PROBABILITY

Example: Candy pmf

- To calculate the pmf for the number of candies we use

$$\Pr\{z\} = \sum_{x \in S_x} \underbrace{\Pr\{z|x\}\,\Pr\{x\}}_{\Pr\{z,x\}},$$

where x is either 'Sara tosses a coin' or 'Sara throws a dice'.

40% 60%

- Hint: First calculate the joint probability of $\Pr\{z, x\}$

$\propto \Pr\{z \mid x=\text{"coin"}\}$

$\propto \Pr\{x \mid z=5\}$

$\Pr\{z, x\}$	z=0	z=1	z=2	z=3	z=4	z=5	z=6	$\Pr\{x\}$
x = coin	0.2	0.2	0	0	0	0	0	0.4
x = dice	0	0.1	0.1	0.1	0.1	0.1	0.1	0.6
$\Pr\{z\}$	0.2	0.3	0.1	0.1	0.1	0.1	0.1	

$\propto \Pr\{z \mid x=\text{"dice"}\}$

A primer in statistics – Expectation, covariance and the Gaussian distribution

Sensor fusion & nonlinear filtering

Lars Hammarstrand

EXPECTED VALUE AND COVARIANCE

- Probability distributions are often characterized by their mean vectors and covariance matrices.

Expected value (mean vector)

- The expected value (mean) of a random vector $\mathbf{x} = [x_1, x_2, \ldots, x_m]^T$ is

μ

$\bar{x}$

$$\mathbb{E}\{\mathbf{x}\} = \int \mathbf{x}\, p(\mathbf{x})\, d\mathbf{x}$$

where $\int d\mathbf{x}$ is shorthand for $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dx_1 \ldots dx_m$.

Covariance matrix

- The covariance matrix is ($\mathbf{x}$ is a column vector)

$$\mathrm{Cov}\{\mathbf{x}\} = \mathbb{E}\left\{\underbrace{[\mathbf{x} - \mathbb{E}\{\mathbf{x}\}]\,[\mathbf{x} - \mathbb{E}\{\mathbf{x}\}]^T}\right\}$$

m

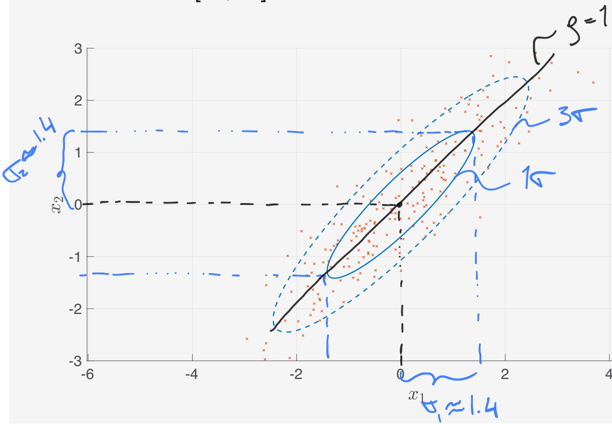
m

- For discrete-valued random variables the above integrals are replaced by the corresponding summations.

GUESS THAT COVARIANCE

Example: Guess that covariance

- Suppose we have independent samples from a zero-mean random vector $\mathbf{x} = [x_1, x_2]^T$. What is the covariance matix of $\mathbf{x}$?



$$\mathrm{Cov}\{x\} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1.4^2 & 0.9\,1.4^2 \\ 0.9\,1.4^2 & 1.4^2 \end{bmatrix}$$

$$\approx \begin{bmatrix} 2 & 1.8 \\ 1.8 & 2 \end{bmatrix}$$

LAW OF LARGE NUMBERS

- The law of large numbers states that sample averages converge to expected values.

Law of large numbers

- If $x_1, x_2, \ldots$ are independent and identically distributed random variables distributed according to $p(x)$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} x_i = \mathbb{E}_{p(x)}\{x\}.$$

Example: Throwing a dice many times...

- ...the average face value converges to the expected value

$$\frac{1+2+3+4+5+6}{6} = 3.5$$

GAUSSIAN DISTRIBUTIONS

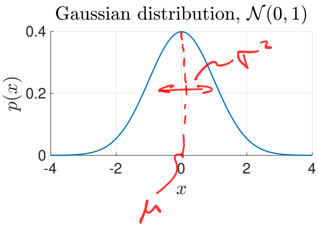
- The most important distribution is the Gaussian distribution (at least in this course).

Gaussian distribution

- We write $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{Q})$ to denote that $\mathbf{x}$ is a Gaussian random variable with mean $\boldsymbol{\mu}$ and covariance $\mathbf{Q}$.
- The pdf of $\mathbf{x}$ is

$$p(\mathbf{x}) = \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \mathbf{Q}) = \frac{1}{\sqrt{|2\pi\mathbf{Q}|}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{Q}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

where $|\cdot|$ denotes the determinant.



- Note: the pdf of a Gaussian random variable is completely determined by its mean and its covariance matrix.

GAUSSIAN DISTRIBUTIONS

Linear combination of indep. Gaussian random variables

- Let $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}_x, \mathbf{Q}_x)$ and $\mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}_y, \mathbf{Q}_y)$.
- Then a linear combination of $\mathbf{x}$ and $\mathbf{y}$,

$$\mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y},$$

where $\mathbf{A}$ and $\mathbf{B}$ are deterministic matrices, is also Gaussian with mean

$$\boldsymbol{\mu}_z = \mathbb{E}\left\{\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y}\right\} = \mathbf{A}\boldsymbol{\mu}_x + \mathbf{B}\boldsymbol{\mu}_y$$

and covariance

$$\begin{aligned}\mathbf{Q}_z &= \mathrm{Cov}\left\{\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y}\right\} = \mathrm{Cov}\left\{\mathbf{A}\mathbf{x}\right\} + \mathrm{Cov}\left\{\mathbf{B}\mathbf{y}\right\} + \underbrace{\mathrm{Cov}\{Ax, By\}}_{=0} + \underbrace{\mathrm{Cov}\{By, Ax\}}_{=0} \\ &= \mathbf{A}\mathbf{Q}_x\mathbf{A}^T + \mathbf{B}\mathbf{Q}_y\mathbf{B}^T.\end{aligned}$$

PROBABILITY THEORY – KEY RESULTS

Conditional distributions:
$$\begin{cases} p(x, z) = p(z|x)p(x) \\ p(z|x) = \frac{p(x,z)}{p(x)} \end{cases}$$

Law of total probability:
$$\begin{cases} p(z) = \int_x p(x, z)\, dx \\ p(z) = \int_x p(z|x)p(x)\, dx \end{cases}$$

1st and 2nd moments:
$$\begin{cases} \mathbb{E}\{\mathbf{x}\} = \int \mathbf{x}\, p(\mathbf{x})\, d\mathbf{x} \\ \mathrm{Cov}\{\mathbf{x}\} = \mathbb{E}\left\{[\mathbf{x} - \mathbb{E}\{\mathbf{x}\}][\mathbf{x} - \mathbb{E}\{\mathbf{x}\}]^T\right\} \end{cases}$$

Gaussian pdf:
$$\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \mathbf{Q}) = \frac{1}{\sqrt{|2\pi\mathbf{Q}|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \mathbf{Q}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right)$$