Self-evaluation

In this course, we assume a certain familiarity with probability theory and linear algebra. You should be able to answer the following questions without too many problems. If you find you have difficulty, please consult the supplied primers or an appropriate textbook.

For probability theory, some good textbooks are:

- C.M. Grinstead and J.L. Snell, *Introduction to probability*, 2^{nd} edition. Providence, RI:American Mathematical Society, 1997. ISBN: 0821807498. Freely available online at `https://math.dartmouth.edu/~prob/prob/prob.pdf`.
- D.P. Bertsekas and J.N. Tsitsiklis, *Introduction to probability*, 2^{nd} edition. Bellmont, MA:Athena Scientific, 2008. ISBN: 9781886529236.
- A. Drake, *Fundamentals of applied probability theory.* New York, NY:McGraw-Hill, 1988. ISBN: 0070178151.
- S. Ross, *A first course in probability.* Upper Saddle River, NJ:Prentice Hall, 2005. ISBN: 0131856626.

For linear algebra:

- J. Hefferson, *Linear algebra.* 2008. Freely available online at `http://joshua.smcvt.edu/linearalgebra/`.
- R. Beezer, *A first course in linear algebra.* 2008. Freely available online at `http://linear.ups.edu/online.html`.
- G. Strang, *Linear algebra and its applications*, 4^{th} edition. Pacific Grove, CA:Brooks Cole, 2005. ISBN: 0030105676.
- D. Poole, *Linear algebra: a modern introduction*, Pacific Grove, CA:Brooks Cole, 2008. ISBN: 0534341748.

Question 1 – Probability and combinatorics

From an ordinary deck of 52 cards, someone draws a card at random.

(a) What is the probability that it is a black ace or a red queen?

(b) What is the probability that it is a face card or a black card?

(c) What is the probability that it is neither a heart nor a queen?

(d) If two random cards were missing from the deck, what is the probability that the next card drawn is a spade?

Question 2 – Expectation, variance, sample estimate

The time it takes for a student to finish a statistics aptitude test has probability density function (PDF)

$$f(x) = \begin{cases} 6(x-1)(2-x) & 1 < x < 2 \\ & \text{otherwise} \end{cases} \tag{1}$$

(a) Determine the expected time it takes for a randomly selected student to finish the test.

(b) Determine the standard deviation of this time.

(c) Given 3 students who finish their test in 1.5 hours, 2 in 1.67 hours, 1 in 1.2 hours and 1 in 1.9 hours, calculate the sample mean μ and standard deviation σ for this sample from $f(x)$.

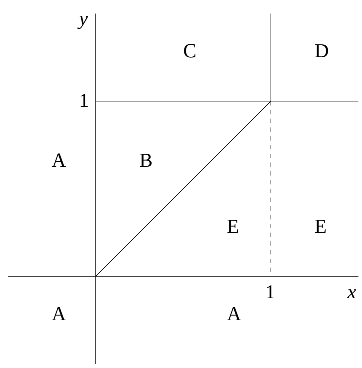
Question 3 – The normal distribution

Suppose that the IQ of students at Delft University is normally distributed with mean 125 and standard deviation 10, respectively. What is the probability that of 20 randomly selected Delft students, five have an IQ of at least 135? Use the on-line calculator at `http://stattrek.com/online-calculator/normal.aspx`, to calculate the value of the standard normal CDF $\Phi(x)$.

Question 4 – Multivariate CDFs and PDFs

The joint probability density function (PDF) of two random variables X and Y is given by:

$$f_{X,Y}(x,y) = \begin{cases} kxy & 0 \le x \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$



(a) Show that in area B (see figure) the expression for the cumulative distribtion function (CDF) is given by $F_{X,Y} = \frac{1}{8}kx^2(2y^2 - x^2)$.

(b) Determine the value of k.

(c) Give expressions for the cumulative distribution function on the other four areas (A, C, D and E).

(d) Calculate the marginal probability density functions of x and y, $f_X(x)$ and $f_Y(y)$.

(e) Find the expected value and variance of Y.

(f) Are the random variables X and Y independent? Motivate.

Question 5 – Correlation, covariance and independence

A stick of length 1 is broken into two pieces at a random point (drawn from a uniform distribution).

(a) Calculate the covariance between the lengths of the two pieces.

(b) Calculate the correlation coefficient between the lengths of the two pieces.

(c) Are the lenghts independent?

Question 6 – The multivariate normal distribution

Let X_1 and X_2 be random variables with a joint multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Writing $\boldsymbol{x} = [x_1, x_2]^T$:

$$f(\boldsymbol{x}) = \frac{1}{2\pi|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$

The covariance matrix can be written as

$$\boldsymbol{\Sigma} = \boldsymbol{DRD}$$

where $\boldsymbol{D}$ is a diagonal matrix with the standard deviations on the diagonal, $D_{ii} = \sigma_i$, and R_{ij} is the correlation coefficient between x_i and x_j. Prove that if X_1 and X_2 are not correlated, they are also independent (note that this only holds for the normal distribution!).

Question 7 – Bayes' rule

Suppose that for the general population, 1 in 1000 people carries the human immunodeficiency virus (HIV). A test for the presence of HIV yields either a positive (+) or negative (-) response. Suppose further that the test gives the correct answer 95% of the time.

(a) Calculate $P(-|H)$, the conditional probability that a person tests negative given that that person does have HIV.

(b) Calculate $P(H|+)$, the conditional probability that a randomly chosen person has HIV given that the test is positive.

(c) In a small town of 100 people, what is the probability that at least one person tests positive for HIV? .