Exam: MVE051/MSG810 – Matematisk statistik och diskret matematik, MSG820 – Statistik för fysiker
Time and place: Thursday January 14, 2016, 14:00-18:00, Hörsalsvägen
Examinator: Marcin Lis, tel. 73 75 54 583
Allowed help: Chalmers approved calculator, Beta handbook, Swedish-English dictionary
Grades: Chalmers: 3: 12 points, 4: 18 points, 5: 24 points. GU: G: 12 points, VG: 21 points. Maximal amount of points is 30

You should always justify your answer except in problem 1, where you do not need to do that.
Lycka till!

1. ($6 \times 0.5 = 3$ points) This quiz tests your understanding of fundamental notions of probability theory. Decide whether the following statements are **true** or **false**. You do not need to justify your answer.

   (a) For any two events A and B such that $\mathbf{P}(A) = \mathbf{P}(B) = 0.9$, it is true that
   $$\mathbf{P}(A \cap B) \geq 0.8$$

   (b) For any two events A and B such that $\mathbf{P}(A) = \mathbf{P}(B) = 0.3$, it is true that
   $$\mathbf{P}(A \cup B) \geq 0.6$$

   (c) For any random variable X, it is true that
   $$\mathbf{E}\left[\frac{1}{2}(X + 2016)\right] \geq \mathbf{E}[X]$$

   (d) For any random variable X, it is true that
   $$\mathrm{Var}\left[\frac{1}{2}(X + 2016)\right] \leq \mathrm{Var}[X]$$

   (e) If $\mathbf{P}(A \mid B) = \mathbf{P}(A)$ for two events A and B such that $\mathbf{P}(B) > 0$, then A and B are independent.

   (f) If $\mathbf{E}[XY] = \mathbf{E}[X]\,\mathbf{E}[Y]$ for two random variables X and Y, then X and Y are independent.

2. ($2 + 2 = 4$ points)

   (a) Compute the expected number of points that an unprepared student gets for solving problem 1 by guessing at random. You can assume that
   - each question is worth 0.5 point,
   - the probability of guessing right for each question is 50%,
   - guesses for different questions are independent of each other.

   (b) Suppose that there are 100 unprepared students who are solving problem 1 independently of each other. Using the central limit theorem, compute the probability that the total number of points that they score is at least 200.

3. ($2+1+1 = 4$ points) Let X be a continuous random variable with the probability density function
   $$f(x) = \begin{cases} (1+x)/2 & \text{for } -1 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

   Compute the following:

   (a) $\mathbf{E}[2X + 1]$,

   (b) $\mathbf{P}(X > 0)$,

   (c) $\mathbf{P}(X = -1/2 \text{ or } X = 1/2)$.

4. $(2+1=3$ points) Let

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

be the transition matrix of a Markov chain on the state space $\{A, B, C, D, E\}$.

(a) Draw the state diagram of the chain. Identify all transient and absorbing states. Is this chain absorbing?

(b) Argue without computations why the probability that the chain started in state C is absorbed in state E is $1/2$.

5. (3 points) Compute the generating function of the sequence $(a_n)_{n=1}^{\infty}$ which is given by the following recursive formula:

$$a_0 = 0, \qquad a_1 = 1, \qquad \text{and} \qquad a_n = 3a_{n-1} + 2a_{n-2}, \qquad \text{for } n \geq 2.$$

You **do not need** to compute the sequence $(a_n)_{n=1}^{\infty}$ itself.

6. $(1+3=4$ points) A doctor claims that the weight of a newborn baby has grown compared to the last year and is now 3.7 kg. He weighs a sample of six babies and gets the following results

$$3.5, \quad 3.6, \quad 4.0, \quad 4.1, \quad 3.8, \quad 4.2.$$

(a) Write down the research and null hypothesis. Describe the assumptions that you have to make concerning the data.

(b) Compute the p-value of the observed results.

7. $(2+3=5$ points) John wants to know how much fuel his car uses on average. He makes five trips and writes down the distance traveled and the amount of fuel used. Here are his results:

distance [km]	100	150	300	350	400
used fuel [l]	5.9	7.6	14.6	17.2	19.8

(a) Write down the assumptions of the linear regression model.

(b) Estimate the amount of fuel used per 100km by performing linear regression.
Hint: Use 100km as the unit of distance.

8. $(2+2=4$ points)

(a) Let $X_1, X_2, X_3, \ldots$ be a sequence of independent random variables with the same dsitribution X ($X_i \sim X$ for $i = 1, 2, 3, \ldots$). Assume that

$$\mathbf{E}[X] = 0 \quad \text{and} \quad \mathrm{Var}[X] = 2016,$$

and let

$$S_n = \frac{X_1 + X_2 + \ldots + X_n}{n}.$$

Using Chebyshev's inequality, find N such that

$$\mathbf{P}(|S_n| \geq 10) \leq 0.01 \qquad \text{for all } n \geq N.$$

(b) Let X be continuous random variable satisfying $\mathbf{E}[X] = 0$. Prove that for all $\varepsilon > 0$,

$$\mathbf{P}(|X| \geq \varepsilon) \leq \frac{\mathbf{E}[X^4]}{\varepsilon^4}.$$

Hint: Modify the proof of Chebyshev's inequality.