Final Exam Probability and Statistics WI2180LR-II
29 October 2018, 13.30 – 15.30

Only the use of a **non-graphical** calculator and a clean copy of the formula sheet is allowed.
This exam consists of nine multiple choice questions and three open questions. You should answer the open questions on the exam sheet.

Grade: Every correct multiple choice question counts for 2 points; for the open questions, points are denoted per part. Then:

$$\text{Grade} = \frac{MC + OQ}{4} + 1,$$

where MC is the amount of points scored for the multiple choice part and OQ for the open questions part.

Explanation MC sheet: Colour the boxes black or blue. **Fill in the version**, course code, your name and **student number**. The latter should be ticked as well. Finally, sign the sheet with your signature.

Multiple choice questions
Version 1

1. Suppose you were to obtain 100 random samples of the same sample size and construct 100 confidence intervals for the parameter θ with confidence level 95%.

   What is the probability that all of them will contain the true value of θ?

   a. 0.9500 **b.** 0.0500 **c.** 0.9900 **d.** 0.010 **e.** 0.0059 **f.** 0.9941

2. Octopus Paul gained worldwide fame in 2010 because he correctly predicted the winner of all seven matches of Germany during the World Cup football. Furthermore, he correctly predicted the World Cup final, which Spain unfortunately won against The Netherlands. So in total Paul predicted 8 matches out of 8 correctly. Assume that Paul correctly predicts the winner of a match with probability p. Furthermore, assume that a match can not end with a draw.

   You wish to investigate whether Paul has exceptional prediction abilities by using a statistical test. What are the relevant hypotheses?

   a. $H_0 : p = 0.5,\ H_1 : p > 0.5$ **b.** $H_0 : p < 1,\ H_1 : p = 1$
   c. $H_0 : p = 1,\ H_1 : p < 1$ **d.** $H_0 : p > 0.5,\ H_1 : p = 0.5$

3. During your well-earned holiday in the south of France you decide to see whether you are lucky in a casino in Monaco. You play the game 'Sept': you roll two dice and if the sum of the spots equals seven, you win 5 euro. In every other case, you lose 1 euro. You decide to play this game 120 times.

   What is approximately the distribution of your total earnings?

   a. $N(-50, 400)$ **b.** $Bin(120, \frac{1}{11})$ **c.** $Bin(120, \frac{1}{2})$
   d. $N(0, 600)$ **e.** $N(0, 400)$ **f.** $Bin(120, \frac{1}{6})$

4. Consider the following sample statistics (numerical summaries):

   $\bar{x}_n$: sample mean m: sample median
   s: sample standard deviation IQR: Interquartile Range.

   Which of these statistics provide information about the location of a dataset?

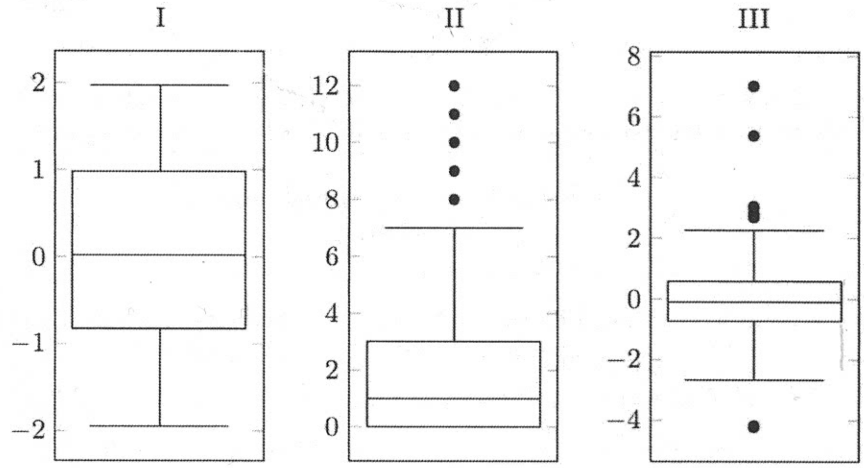
   a. s and IQR **b.** $\bar{x}_n$ and m **c.** m and IQR
   d. $\bar{x}_n$ and IQR **e.** m and s **f.** $\bar{x}_n$ and s

5. You wish to estimate the parameter $-1 \leq \theta \leq 4$ using the maximum likelihood principle. Based on the data, you find the likelihood function $L(\theta) = \theta^3 - 5\theta^2 + 3\theta + 9$.

   Determine the maximum likelihood estimate of θ.

   a. 3 b. -1 c. $\frac{1}{3}$ d. $\frac{5}{2}$ e. $\frac{5}{3}$ f. $-\frac{1}{2}$

6. Three datasets, each of sample size 200, have been sampled from three (not necessarily different) distrubutions. Consider the boxplots of these datasets below.



   From which distributions have these datasets been sampled?

   a. I: Normal, II: Geometric, III: Normal
   b. I: Uniform, II: Geometric, III: Normal
   c. I: Uniform, II: Geometric, III: t-distribution
   d. I: Normal, II: Geometric, III: t-distribution
   e. I: Uniform, II: Exponential, III: Normal
   f. I: Normal, II: Exponential, III: Normal
   g. I: Uniform, II: Exponential, III: t-distribution
   h. I: Normal, II: Exponential, III: t-distribution

7. Let $X_1, \ldots, X_{100}$ be i.i.d. random variables, having a *Par*(3) distribution. Define $S = \sum_{i=1}^{100} X_i$. Furthermore, let $Z \sim N(0,1)$.

   For which value of a does it hold that $P(S \geq 180) \approx P(Z \geq a)$?

   a. 0.35 b. 3.46 c. -1.39 d. 4 e. -1.6 f. None.

8. Let X have an exponential distribution with parameter λ. You wish to test $H_0 : \lambda = 2$ against $H_1 : \lambda > 2$. You reject H_0 in favour of H_1 if $X \leq 0.25$.

   Calculate the probability of a type I error. Round to two decimal places.

   a. 0.88 b. 0.22 c. 0.12 d. 0.39 e. 0.78 f. 0.61

9. Let X have a *Pois*(μ) distribution. Define $p = P(|X - \mu| \geq 2\mu)$. Which of the following statements is the strongest statement[1] that can be deduced using Chebyshev's inequality?

   a. $p \geq \frac{1}{2}$ b. $p \geq \frac{1}{4\mu}$ c. $p \leq \frac{1}{2\mu}$ d. $p \leq \frac{1}{4\mu}$
   e. $p \leq \frac{1}{4}$ f. $p \leq \frac{1}{2}$ g. $p \geq \frac{1}{4}$ h. $p \geq \frac{1}{2\mu}$

[1] *Strongest statement* means that if you e.g. deduced that $p \leq 0.40$, then $p \leq 0.60$ is also true, but $p \leq 0.40$ is a stronger statement. Similarly, if you deduced that $p \geq 0.60$, then $p \geq 0.40$ is also true but $p \geq 0.60$ is stronger.