Sample Final

1 Chapter 1

1) (1 point) Assume you have a population of 100 items, with item numbers 00, 01, 02, ... 99. You choose a sample of size 10 by choosing a random number x between 0 and 9, and selecting all items whose number ends with the digit x. For instance, you may select the fifth, fifteenth, 25th, 35th, 45th, 55th, 65th, 75th, 85th, and 95th item in a sample if $x = 5$.

Do the samples obtained in this way obey the properties for a random sample? Why or why not?

2 Chapter 2

2) (3 points) The data on sheet 1 describes the age of the 64 employes of a small company. Make a histogram, graphing the data using the 12 classes 18-22, 23-27, 28-32, 33-37, 38-42, 43-47, 48-52, 53-57, 58-62, 63-67.

3) (3 points) In a set of 10 scores the value 2 occurs three times, the value 4 occurs twice, 6 occurs twice, and 7 occurs three times. What is the mean of the scores?

3 Chapter 3

4) (3 points) What are Mean and Median of the x-data of the it sample in Data set on sheet 3. What is the standard deviation.

5) (Chebyshev) The mean salary of 1000 workers of a company is 33000 Euro (per year), the standard deviation is 3000 Euro. Use Chebyshev's Theorem to make a statement about the number of employees with salaries between 27000 and 39000 DM.

4 Chapter 4

6) (3 points) A biased die has the following probability distribution for number of dots observed on the upward face:

dots	probability
1:	0.2
2:	0.15
3:	0.15
4:	0.15
5:	0.2
6:	0.15

a) What is the expected number of dots shown?

b) What is the standard deviation for that random variable?

c) Rolling the die 1000 times and adding the dot numbers, would you expect to get a larger sum with this biased die or with a unbiased die?

7) (Chapter 4, 2 points) One card is randomly selected from a 52 card deck. Find the probability that the selected card is

a) an ace,

b) a black card,

c) an ace or a black card,

e) an ace or a king or a queen.

8) (Chapter 4, 3 points) A geologist reports a 65% chance of finding oil at a certain site. The company starts drilling. During the drilling, sample cores are takenfrom the well and studied. These sample cores have a history of predicting oil when there is oil in about 85% of the time. However, about 6$ of the time the sample cores will predict oil when there is no oil. Our geologist is delighted, since the sample cores predict oil for this well.

Use the new information from the sample cores to revise the geologist's original probability the well will hit oil. What is the new probability?

5 Chapter 5

9) (Chapter 5, 2 points) Desribe the features of a Binomial experiment.

10) (Chapter 5, 3 points) Sociologists say that 90% of all married women dislike their mother in law. Suppose that six married women are having coffee together. What is the probability that

- (a) all of them dislike their mother-in-law,

- (b) none of them dislike their mother-in-law,

- (c) at least four of them dislike their mother-in-law,

11) (Chapter 5, 3 points) The probability that a single proof reader will detect a typo in a manuscript is 0.85.

- (a) How many proof reader are required to be 99% sure that any given typo is found.

- (b) Assume there are four proof reader, and originally the manuscript contains 100 typos. How likely is it that these proof reader together detect at least 96 of these typos? What is the expected value of typos detected?

6 Chapter 6

12) (Chapter 5, 2 points) What is a normal distribution? What is a standard normal distribution?

13) (2 points) The continuous random variable x has a normal distribution with mean $\mu = 30$ and standard deviation $\sigma = 4$.
    a) Compute the probability $P(x \geq 38)$ that x is larger or equal to 38.
    b) Compute the probability $P(24 \leq x \leq 26)$ that x lies between 24 and 26.

7 Chapter 7

14) (5 points) Explain the notions sample, population, parameter, statistics, random variable.

15) (2 points) What is a sampling distribution?

16) (2 points) Suppose that x has a distribution with $\mu = 25$ and $\sigma = 3.5$.

    a) If a random samples of size $n = 9$ are selected, can we say anything about the $\overline{x}$-distribution of sample means?

    b) If the original x-distribution is normal, can we say anything about the $\overline{x}$-distribution from samples of size 9? Find $P(23 \leq \overline{x}le26)$.

17) (2 points) *All* possible samples of size 4 are selected from a population of 30, and the mean of each sample is determined.

    a) How many different samples of size 4 can we choose?

    b)) The mean of the sample means will be

        i) exactly the same as the population mean.
        ii) larger than the population mean.
        iii) smaller than the population mean.
        iv) cannot be estimated in advance.
        v) none of the above.

18) (Chapter 7, 2 points) Formulate the Central Limit Theorem

8 Chapter 8

19) (Chapter 8, $3 + 2 + 1$ points) Assume that a variable x has a normal distribution in a population.

   a) Construct a 99 % confidence interval estimation for μ based upon the data given on Sheet 4.

   b) Would the confidence interval change if we would know that the population standard deviation σ equals 14? If it would, how would it change?

   c) Would the 95% confidence interval for (a) be larger or smaller?

20) (Estimating) Assume $5 \leq \mu \leq 9$ is the 95 % confidence interval estimate for the population mean μ, based on a sample of 40 items. Which of the following is true?

   i) the population mean μ must lie between 5 and 9.

   ii) the sample mean $\overline{x}$ equals 7.

   iii) the probability that the population mean lies between 5 and 9 is 5 %.

   iv) the probability that the population mean lies between 5 and 9 is 95 %.

   v) the probability that the population mean lies between 5 and 9 is 97.5 %.

21) (3 points) A sample of 200 people were asked to identify their major source of news information. 110 stated that their major source was television news coverage. Construct a 95 % confidence interval of people in the population who consider television their major source of news information.

9 Chapter 10

22) (Scatter diagram, correlation coefficient)
   a) Develop a scatter diagram for age versus time in data set sheet 2.
   b) Find the correlation coefficient for the data.
   c) Is there some relation between x and y? Of what kind?
   d) Compute the regression line of time, depending on age. What time would you expect for somebody born in 1940? Use Chebyshevs Theorem to give an interval for the possible time for such a person with probability 75%.

23) (1+2 points) How is the regression line for paired data defined?

24) (1+2 points) Describe what the stadard error of estimation for paired data describes.