#### Voorblad voor een schriftelijk tentamen/toets

##### Verplicht gedeelte:

Vaknaam : Statistics 1 for IBA

Vakcode : 30J109

Datum tentamen : 25 maart 2011

Duur tentamen : 3 uur

Docent : Gert Nieuwenhuis (tel. 2350 of 0650504640)

Telefoon secretariaat departement: 2430

Studenten worden geacht zich tijdens het tentamen correct te gedragen en de instructies van examinator en surveillant op te volgen.

**Bij constatering van fraude wordt streng opgetreden.**

**Alle studenten moeten uiteindelijk inleveren:**

1. de door hen geproduceerde oplossingen van de opgaven;
2. ook de formulieren met alle vragen.

**Vriendelijk verzoek aan de surveillanten**: zou u de uitwerkingen van de studenten per blauwe lijst bij elkaar willen leggen? Zodat onder elke blauwe lijst meteen de toetsen liggen van de studenten die op deze lijst staan. Wel graag de formulieren met de opgaven apart op een stapel. Alvast zeer bedankt.

***Facultatief gedeelte:***

Open boek examen neen

Schrapkaart neen

Gebruik eenvoudige calculator ja

Gebruik grafische calculator neen

Enkelzijdig afdrukken nee

Kladpapier beschikbaar ja

Uitwerkingen dienen op tentamenpapier gemaakt te worden, iedereen dient zijn werk **en** de formulieren met de vragen in te leveren.

Mid-Exam of Statistics 1 for IBA

25 March 2011; time: 3 hours

It is **NOT** allowed to use a graphical, programmable calculator; only a simple pocket calculator is allowed. **Always explain your answers**.

This exam contains 6 exercises on 3 pages.

**At the end, you have to hand in**:

1. your solutions to the six exercises
2. these original sheets with the six exercises

On Monday, March 28, these six exercises and their solutions will be posted on Blackboard.

**Exercise 1**

For a sample of 500 people, randomly drawn from the population of Dutch adults applying to get a driver's license, the variable X = ‘number of times the person needs to take the test to receive the driver's license’ is measured. As a result, the following frequency distribution is obtained:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **value** | 1 | 2 | 3 | 4 | 5 | 6 | **total** |
| **frequency** | 220 | 123 | 104 | 39 | 8 | 6 | 500 |

1. Determine the relative frequency of people in the sample who needed to take the test more than 3 times.
2. Let F be the accompanying cumulative distribution function. Calculate F(3.2).
3. Calculate the mean of the 500 observations in the sample.
4. Also calculate the accompanying variance.
5. What is the median of this sample dataset?

**Exercise 2**

You want to explain how the variable *W* = 'daily number of hours that you study'

varies with the variable *V* = 'daily number of hours that you watch television' by

doing a simple regression of *W* on *V* on the basis of a random sample of 7 days.

For these days, you have the following observations of *W* and *V*:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **day *i*** | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| ***wi*** | 4 | 2 | 0 | 2 | 2 | 0 | 1 |
| **vi** | 3 | 5 | 6 | 4 | 3 | 7 | 6 |

Below, you can freely use (so, without proving) that:

; ; ;

1. Calculate the covariance and the correlation coefficient of the v- and the w-data. What do you think of the (strength of the) linear relationship of W on V? Why?
2. Determine the equation of the sample regression line of w on v.

**Exercise 3**

Consider a random experiment with sample space . With respect to the events A, B, D1, D2, and D3, the following is given:

*D*1, *D*2, and *D*3 form a partition of the sample space 

B and D1 are independent; B and D2 are disjoint;

; ; ; ;

1. Calculate .
2. Calculate .
3. Calculate .
4. Can *B* and be disjoint? Why (not)?
5. Can and be independent? Why (not)?
6. Calculate .
7. Calculate .

**Exercise 4**

Suppose the winning number in a lottery is a four-digit number determined by randomly drawing four slips of paper (without replacement) from a box that contains nine slips numbered consecutively 1 through 9 and then recording the four digits in order from smallest to largest.

1. How many different lottery numbers are possible?
2. Find the probability that the winning number has only odd digits.
3. How many different lottery numbers are possible if the digits are recorded in the order they were drawn?

**Exercise 5**

Mr. Bean has become the CEO of a major company doing imports and exports of various types of beans around the world. To model next year’s profit X (in millions of euro), he decided to make use of the following probability density function (pdf) f for X :

1. Does this model assume that X is discrete or X is continuous? How do we know that?
2. Determine the accompanying cumulative distribution function (pdf) F(x) for each of the following situations:

; ; ; ;

1. Calculate and .
2. Calculate the expectation μ and the standard deviation σ of X.
3. Calculate and compare the answer with the result that follows from Chebyshev’s rule.

Next, consider the random variable (in millions of euro).

1. Calculate the expectation and standard deviation of *Y*.
2. Calculate the probability that *Y* will be larger than *X*.

**Exercise 6**

Ms. Starducks, the owner of a chic restaurant, knows that the weekly profit is usually between 2000 and 6000 euro, with an average of 4000 euro per week. To model next week's profit X (in units of 1000 euro) she decided to use a triangular probability density function, as follows:

Here, *a* is a constant. Note that the graph of *f* is as follows:

2

4

6

*a*

*f*(*x*)

1. Carefully explain why the constant *a* has to be equal to 0.5.
2. What is the median of *X*? Why?
3. Ms. Starducks labels a week with less than €2800 profit as 'miserable'. Calculate the probability that next week will be a 'miserable' week.
4. She also labels a week with more than €5000 profit as a 'great' week. Determine the percentage of 'great' weeks.
5. Calculate the conditional probability .
6. Find the 12.5%-percentile of X; that is, calculate the real number b for which .