Mathematical Statistics, Final Exam, WNE UW, June 2009

Fill in the gaps [".........."]. 1 question (•) = 1 point; maximum = 12 points. Write down your answer **after you have checked it**; unreadable answers will be evaluated as **wrong**! You can use a calculator, statistical tables, notes and/or textbooks. **Do not communicate; do not use mobile phones or computers; do not cheat**!

NAME: ....

1. Let $X_1, X_2, \ldots, X_n$ be *i.i.d.* random variables with probability density given by

$$f_\theta(x) = \begin{cases} \theta x^{-\theta-1} & \text{for } x \geq 1; \\ 0 & \text{for } x < 1 \end{cases}$$

(a *Pareto* distribution), where $\theta > 0$ is an unknown parameter.

- Compute the *maximum likelihood estimator* (MLE) of parameter θ, given the sample $X_1, X_2, \ldots, X_n$:

  $\hat{\theta}_{\mathrm{ML}} =$ ....

- Compute the estimator of θ by the *method of moments* (MME):

  $\hat{\theta}_{\mathrm{MM}} =$ ....

2. A survey was conducted to examine the relation between the level of education and tolerance. There were 300 people in the survey and the results were the following:

	tolerance	lack of tolerance	total
university graduate	60	40	100
high school graduate	50	50	100
no high school	40	60	100
total	150	150	300

- Compute the *test statistic* χ^2 to test for independence between the level of education and tolerance:

  $\chi^2 =$ ....

- Compute the $p - value$ and interpret the result:

  $p =$ ...., therefore we reject/do not reject the null hypothesis that the row and column variables are independent of each other (mark the right answer).

Hint: The $\chi^2(2)$ distribution (chi-square with 2 degrees of freedom) is the exponential distribution $\mathrm{Ex}(1/2)$.

3. 10 packets of butter have been weighed. The outcome of the measurement is listed below:

   $$250;\ 240;\ 243;\ 247;\ 248;\ 249;\ 251;\ 242;\ 246;\ 244.$$

   Assume this is an *iid* sample from $N(\mu, \sigma^2)$, with μ and σ unknown.

   - Give a confidence interval for the average weight of one packet μ at the confidence level $1 - \alpha = 0.95$.

     [...................; ....].

   - Test the null hypothesis $H_0 : \sigma \leq 5$ against the alternative hypothesis $H_1 : \sigma > 5$. Give the value of the test statistic, compare with the 0.95-quantile of the appropriate χ^2 distribution and make decision.

     statistic = ... Since the critical region is the interval ...., we decide to reject/do not reject the null hypothesis (mark the right answer).

4. Our quantity of interest is the percentage of graduates in Warsaw, who would get a job within 6 month after they completed their studies. A sample of graduates has been surveyed and it turned out S out of 400 got a job.

   - Test the null hypothesis H_0 that at least 60% of graduates get a job against the alternative that less than 60% of graduates get a job. The test at significance level $\alpha = 0.05$ *rejects* H_0 if S satisfies the following inequality:

     S..

   - If we observe the value $S = 200$, a confidence interval at the confidence level $1 - \alpha = 0.95$ for the percentage of interest amounts to

     [...................; ....]

   Hint: Use the simplest method and approximate the binomial distribution by a normal distribution.

5. Let $X_1, \dots, X_n$ be an *iid* sample from a uniform distribution $\mathrm{U}(0, \theta)$ with the probability density function

$$f_\theta(x) = \begin{cases} \frac{1}{\theta} & \text{for } 0 \le x \le \theta; \\ 0 & \text{otherwise,} \end{cases}$$

where $\theta > 0$ is an unknown parameter. Consider estimators which are multiples of the sample mean: $\hat{\theta} = c\bar{X}$.

   - Choose constant c such that $\hat{\theta}$ is an *unbiased estimator* of θ:

     $\hat{\theta} =$ ....$\bar{X}$

   - Compute the *variance* of the unbiased estimator obtained above.

     $\mathrm{Var}\hat{\theta} =$ ....

   Remark: We do not recommend to use this estimator because eg. MLE is much better in this model.

6. Let $X_1, X_2, \dots, X_n$ and $Y_1, Y_2, \dots, Y_m$ be two independent samples from the same normal distribution $\mathrm{N}(\mu, 1)$. Two statisticians indpendently perform the test of $H_0 : \mu = 0$ against the alternative $H_1 : \mu > 0$. Both statisticians apply the most powerful test at the significance level $\alpha = 0,05$. One of them uses the sample of Xs whilst the other uses the sample of Ys.

   In fact, H_0 is true (although our statisticians cannot know this).

   - Compute the probability of the event that *both* statisticians will reject H_0: ....

   - Compute the probability of the event that *at least one* of the statisticians will reject H_0: ....