General exam in “Statistical Methods” 16.2.2024

NB! Students are allowed to use a pocket calculator when completing the exam!

1. Define following concepts (briefly):
   a) the central limit theorem (i.e. why are errors usually Gaussian) [2p]
   b) systematic uncertainties (define difference w.r.t. statistical uncert.) [2p]
   c) Fisher’s linear discriminant (application and discriminating function) [2p]
2. Describe the maximum likelihood (ML) method (“suurimman uskottavuuden menetelmä”). Describe also how an ML estimator and its variance are determined. Give the advantages and disadvantages of the ML method. [6p]
3. a) i) During a particular meteor shower, meteors fall at the rate of four per hour. What’s the probability to observe none in half an hour? Estimate also the probability of observing at least one within 20 minutes? [1.5p]

   (ii) The weights of the eggs produced by a farmer’s hens have a standard deviation of 9 g. He feeds some hens a vitamin supplement that will be cost-effective if the average egg weight increases by at least 3.5 g. He obtains 16 eggs from the vitamin-fed hens that have an average weight 4.5 g larger than the eggs from the hens that have not received extra vitamin. Can the farmer conclude that the increase is significant and that he can start feeding all his hens with the supplement? [1.5p]

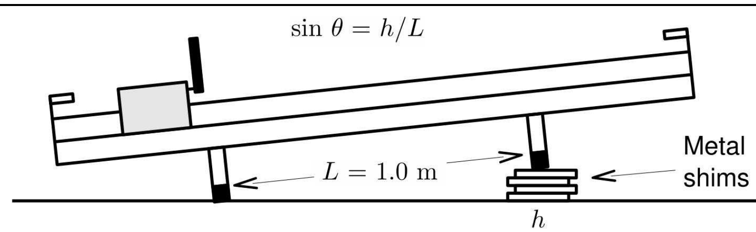
   b) Charged particles traversing a gas volume produce ionization with a mean amount depending on particle type. Suppose a test statistic t based on such measurements follows a Gaussian centered at 0(6) for pions(kaons), both having a standard deviation of 2. Kaon candidates selected by $t > 4$.

   (i) What is the kaon selection efficiency when requiring $t > 4$? [1p]

   (ii) What is the probability that a pion will be accepted as a kaon? [1p]

   (iii) Suppose a sample of particles is known to consist of 95 % pions and 5 % kaons. What is the purity of the kaon sample selected by $t > 4$? [1p]
4. A physics student studied carts sliding down an incline with a negligable friction (an ”air track”, see figure). The data are the measured times for a cart released from rest to travel various distances d along the track. Assume d to be precise and the uncertainty on t to be Gaussian with a σ of 0.05 s.

t [s]	d [cm]	t [s]	d [cm]	t [s]	d [cm]
0.75	1	2.45	16	4.15	49
6.10	100	7.85	144	9.10	225



   (i) Consider the hypothesis $t = \alpha\sqrt{d}$. Find the least square estimator for

α. What is the minimum χ^2 value? Can the hypothesis be excluded? [3p]

(ii) Estimate the uncertainty on α assuming variance of α to be equal to the RCF bound ($= (1 + (\partial b/\partial\alpha))^2 / E\left[-(\partial^2 \ln L/\partial\alpha^2)\right]$, where $b =$ bias). [2p]

(iii) Assume no friction and $g = 9.81$ m/s^2. Estimate inclination angle θ. [1p]

5. In a classic experiment on radioactivity, Rutherford and Geiger counted the number of alpha (α) decays. The table below gives the number of times n_m that they observed m decays of α's in a fixed time interval of 7.5 s. Assuming the source to consist of very many radioactive atoms and the probability for any one of them to emit an α in a short interval is small, one expects the number of decays in a time interval to follow a Poisson distribution. Deviations from this hypothesis would indicate that the decays were not independent e.g. that the emission of an α particle might cause neighboring atoms to decay, resulting in many decays in a short time interval.

m	n_m	m	n_m	m	n_m	m	n_m
0	57	4	532	8	45	12	0
1	203	5	408	9	27	13	1
2	383	6	273	10	10	14	1
3	525	7	139	11	4	> 14	0

(i) Verify with a test statistic whether data are Poisson distributed. [3p]

(ii) What is (approximately) the quality of the Poisson hypothesis. Can the Poisson hypothesis be excluded? [3p]

6. (i) Your professor is a basketball fan but not such a good basketball shooter. Determine analytically his scoring efficiency assuming that he throws the ball randomly and uniformly within a 40 cm distance of the centre of the rim (the "ring" which the ball has to go through to score in basketball). Hint: The uniform distribution is for the ball's geometrical centre. Take care that the whole ball fits into the rim. The circumference of a basketball is 749 mm and the inner diameter of the rim 450 mm. Assume both the ball and the rim to be perfect circles. [3p]
(ii) Repeat (i) taking a Gaussian-distributed and rim-centered throwing probability with a 60 cm width in both projections. Describe in detail how to determine the scoring efficiency using Monte Carlo methods [3p]

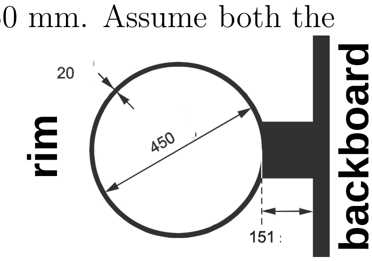


Table of values of the confidence level for different values for the inverse of the cumulative gaussian distribution:

central		one-sided	
$\Phi^{-1}(1-\gamma/2)$	$1-\gamma$	$\Phi^{-1}(1-\alpha)$	$1-\alpha$
1	0.6827	1	0.8413
2	0.9544	2	0.9772
3	0.9973	3	0.9987