NTNU
Norwegian University of Science and Technology

Department of Mathematical Sciences

Examination paper for
Hints for exercise 4 in TMA4295 Statistical inference

Academic contact during examination: Professor Gunnar Taraldsen

Phone: see Wikipage

Examination date: Wednesday in week 36 to Monday in week 38, 2023

Examination time (from–to): Wed 21:00 – Mon 12:00

Permitted examination support material: C

— Tabeller og formler i statistikk, Akademika
– Mathematische Formelsamlung (Matematisk formelsamling) by K. Rottmann
— Stamped yellow A5 sheet with your own handwritten notes
— A specific basic calculator.

Other information:

You may write in English or Norwegian.

All answers must be justified. The answers must include enough details to see how they have been obtained. You must, as always, formulate necessary assumptions as part of the proof of a claim.

All 14 sub-problems carry the same weight for grading.

Language: English

Number of pages: 3

Number of pages enclosed: 0
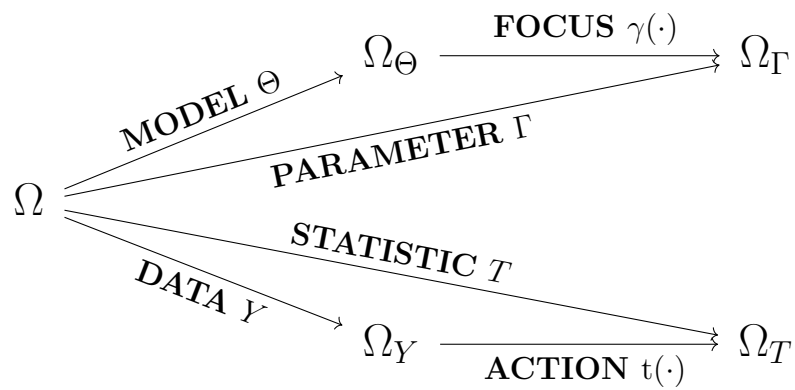
Checked by:

_____
Date   Signature

Figure 1: Functions in statistical inference.

Figure 1 can be useful for the more theoretical part of this exercis. The problem description and tasks are given on the next two pages.

Figure 2: A coiled paper dartboard with 10 beds defined by concentric rings.

Problem 1   Foundations of Statistics

a) Write down and use the definition of the expected value of a simple random variable. Use (ii) as the induction step in a proof by induction.

b) Prove $E_W \phi = \sum_w f(w)\phi(w)$ using the definition of the expectation of simple variables. This is useful by itself!

c) A statistic is a function of the data! An equivalence relation is equivalent with a partition - always. Level sets give a partition. Formulate the sufficiency principle, and use this.

Problem 2   Statistics and Darts

a) The sets in the partition are defined by set differences of disks.

b) Write down the axioms for a family of events and for a probability. Use this. Any event is determined by a sequence of 0's and 1' corresponding to if a set from the partition is not or is included in the event.

c) Write down all ingredients needed for the definition of a statistical model, and verify that this is fine for the example.

d) Use that the score is a simple random variable.

e) This corresponds to the case of throwing a dart n times instead of just one time as in the previous model. It may help to count the number of sets in the resulting partition of the new data space.

f) What is the natural generalization of the binomial distribution?

g) Start by finding an expression for the likelihood.

h) How are the events defined related to the definition of a random variable and its CDF? How is an exponential family defined, and is this relevant here?

i) Explain and use $E(E(X \mid Y)) = E(X)$. Explain that $\text{Var}\, X \geq \text{Var}(X \mid Y)$ and use this. Sufficiency is an important part of the needed argument. Simulation on a computer for calculation of standard uncertainty?

j) Use the definitions.

k) Both the gamma and the chi squared distributions give possible routes for solution.