Preliminary (take home) exam, given Friday 3/31, due Friday 4/7

Exercise P1. (20%) Consider the Gaussian location-scale model (Model III), for sample size n, i. e. observations are i.i.d $X_1, \ldots, X_n$ with distribution $N(\mu, \sigma^2)$ where $\mu \in \mathbb{R}$ and $\sigma^2 > 0$ are unknown. For a certain $\sigma_0^2 > 0$, consider hypotheses $H : \sigma^2 \leq \sigma_0^2$ vs. $K : \sigma^2 > \sigma_0^2$.

Find an α-test with rejection region of form (c, ∞) (i.e. a one-sided test) where c is a quantile of a χ^2-distribution. (Note: it is not asked to find the LR test; but the test should observe level α, on *all* parameters in the hypothesis $H : \sigma^2 \leq \sigma_0^2$.)

Hint: A good estimator of σ^2 might be a starting point.

Exercise P2 (Two sample problem, F-test for variances). Let $X_1, \ldots, X_n$ be independent $N(\mu_1, \sigma_1^2)$ and $Y_1, \ldots, Y_n$ be independent $N(\mu_2, \sigma_2^2)$, also independent of $X_1, \ldots, X_n$ ($n > 1$) where μ_1, σ_1^2 and μ_2, σ_2^2 are all unknown. Define the statistics

$$
\begin{aligned}
F &= F(X, Y) = \frac{S_X^2}{S_Y^2}, \qquad (1) \\
S_X^2 &= n^{-1} \sum_{i=1}^{n} \left(X_i - \bar{X}_n \right)^2, \; S_Y^2 = n^{-1} \sum_{i=1}^{n} \left(Y_i - \bar{Y}_n \right)^2
\end{aligned}
$$

(here (X, Y) symbolizes the total sample).

Define the **F-distribution with k_1, k_2 degrees of freedom** (denoted F_{k_1, k_2}) as the distribution of Z_1/Z_2 where Z_i are independent r.v.'s having χ^2-distributions of k_1 and k_2 degrees of freedom, respectively.

i) (15%) Show that $F(X, Y)$ has an F-distribution if $\sigma_1^2 = \sigma_2^2$, and find the degrees of freedom.

ii) (20%) For hypotheses $H : \sigma_1^2 \leq \sigma_2^2$ vs. $K : \sigma_1^2 > \sigma_2^2$, find an α-test with rejection region of form (c, ∞) (i.e. a one-sided test) where c is a quantile of an F-distribution. (Note: it is not asked to find the LR test; but the test should observe level α, on *all* parameters in the hypothesis $H : \sigma_1^2 \leq \sigma_2^2$).

Exercise P3 (25 %) **(Two sample t-test).** Let $X_1, \ldots, X_n$ be independent $N(\mu_1, \sigma^2)$ and $Y_1, \ldots, Y_n$ be independent $N(\mu_2, \sigma^2)$, also independent of $X_1, \ldots, X_n$ ($n > 1$), where μ_1, μ_2, and σ^2 are all unknown. Consider hypotheses $H : \mu_1 = \mu_2$ vs. $K : \mu_1 \neq \mu_2$.

Show that the likelihood ratio test is equivalent to a certain t-test, i.e. a test where the critical value is chosen as the upper $\alpha/2$-quantile of a t-distribution. (*Equivalence* of two tests based on the same sample here means that they result in the same decisions, for all values of the sample).

Hint: Cp. also homework exercise H6.1. The likelihood ratio computation is similar to proposition 7.5 (ii), p. 81-82 handout.

Exercise P4. (10%) *Complete the proof of proposition 7.5 handout by establishing claim (i).* In detail: consider the Gaussian location-scale model (Model III), for sample size n, i. e. observations are i. i. d. $X_1, \ldots, X_n$ with distribution $N(\mu, \sigma^2)$ where $\mu \in \mathbb{R}$ and $\sigma^2 > 0$ are unknown. Consider hypotheses $H : \mu \leq \mu_0$ vs. $K : \mu > \mu_0$, and the one sided t-test which rejects when the t-statistic

$$T_{\mu_0}(X) = \frac{\left(\bar{X}_n - \mu_0\right) n^{1/2}}{\hat{S}_n}.$$

is too large (with a proper choice of critical value, such that an α-test results). Show that the one sided t-test is the likelihood ratio test for this problem.

Hints: a) When maximizing $p_{\mu,\sigma^2}(x)$ over the alternative, the supremum is not attained ($\mu > \mu_0$ is an open interval). However the supremum is the same as the maximum over $\mu \geq \mu_0$ which is attained by certain maximum likelihod estimators $\hat{\mu}_1, \hat{\sigma}_1^2$ (find these, and also MLE's $\hat{\mu}_0, \hat{\sigma}_0^2$ under H)

b) Note that this time, in difference to part (ii), we have to show that the likelihood ratio is a monotone increasing function of $T_{\mu_0}(X)$ itself, not of its absolute value. Establish this by considering separately the cases of positive and nonnegative values of the t-statistic $T_{\mu_0}(X)$.

Exercise P5 (F-test for equality of variances). Consider the two sample problem of exercise P2, but hypotheses $H : \sigma_1^2 = \sigma_2^2$ vs. $K : \sigma_1^2 \neq \sigma_2^2$.

i) (5%) Find the likelihood ratio test and show that it is equivalent to a test which rejects if the F-statistic (1) is outside a certain interval of form $[c^{-1}, c]$.

ii) (5%) Show that the c of i) can be chosen as the upper $\alpha/2$ quantile of the distribution $F_{r,r}$ for a a certain $r > 0$.