ExerciseFisher: Introduction

In this exercise, the short notation $\text{model}_{K_1,K_2,K_3}$ refers to a DCM model having orders K_1, K_2 and K_3. The number of free parameters η is

$$\eta = K_1 \times (2K_2 + 1) + K_3 + 1$$

The Fisher-test null hypothesis in **Paper I** is

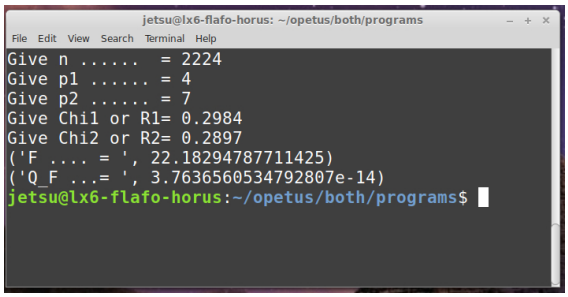
> H_0: *"The* **complex** *model does not provide a significantly better fit to the data than the* **simple** *model"*

This test was applied to Algol's O-C data in **Paper II**. The sample size was $n = 2224$ (All data = **file1** = **1hjdAlgol.dat**). Here, we explain the Fisher-test analysis results given in this extract from Table A8 of **Paper II**.

Model	$\text{model}_{2,1,0}$	$\text{model}_{3,1,0}$	$\text{model}_{4,1,0}$
$\text{model}_{1,1,0}$	↑	↑	↑
$\eta = 4$	$F = 22$	$F = 14$	$F = 11$
$R = 0.2984$	$Q_F = 3.8 \times 10^{-14}$	$Q_F = 2.2 \times 10^{-16}$	$Q_F < 10^{-16}$
$\text{model}_{2,1,0}$	–	↑	↑
$\eta = 7$	–	$F = 6.9$	$F = 5.9$
$R = 0.2897$	–	$Q_F = 0.00012$	$Q_F = 3.6 \times 10^{-6}$
$\text{model}_{3,1,0}$	–	–	←
$\eta = 10$	–	–	$F = 4.9$
$R = 0.2870$	–	–	$Q_F = 0.0021$
$\text{model}_{4,1,0}$	–	–	–
$\eta = 13$	–	–	–
$R = 0.2851$	–	–	–

In this analysis, the sum of squared residuals R (**Paper I**: Eq. 9) was used to compute the Fisher-test test-statistic $F = F_R$ (**Paper I**: Eq. 13).

The input for **fisher.py** was
n = **n** = sample size
η_1 = **p1** = number of free parameters of **simple** model
η_2 = **p2** = number of free parameters of **complex** model
R_1 = **R1** = sum of squared residuals of **simple** model
R_2 = **R2** = sum of squared residuals of **compex** model
Comparison between **simple** model$_{1,1,2}$ and **complex** model$_{2,1,2}$ gave the results shown in this screenshot



The high value $F = 22$ was so unprobable that the critical level reached 3.8×10^{-14}. The H_0 hypothesis **was not rejected**, because

$$Q_F < \gamma_F = 0.001.$$

This meant that the **complex** model$_{2,1,0}$ was better than **simple** model$_{1,1,0}$. The arrow $\uparrow$ pointing towards the **complex** model$_{2,1,0}$ was used to indicate this result. The next four model comparison also gave the same result, because the criterion $Q_F < \gamma = 0.001$ was fulfilled. The first case, where the **simple** model was better than the **complex** model, occured at the lowest line, where the criterion $Q_F < 0.001$ was not fulfilled, and H_0 hypothesis **was rejected**. This is the first case having an arrow pointing to the left $\leftarrow$ towards **simple** model$_{3,1,0}$. In other words, the **simple** model$_{3,1,0}$ was a better than the **complex** model$_{4,1,0}$. There was no longer any reason to test more complex models, and therefore model$_{3,1,0}$ was the best model for the data.

ExerciseFisher: Problem

Print this whole third page and **copy fisher.py** from home-page.

Analysis of another variable star gives the results shown in table below

	$\text{model}_{2,1,2}$ $\eta_2 = \lvert\ldots\ldots\rvert$ $R_2 = 0.05800$	$\text{model}_{3,1,2}$ $\eta_2 = \lvert\ldots\ldots\rvert$ $R_2 = 0.05434$	$\text{model}_{4,1,2}$ $\eta_2 = \lvert\ldots\ldots\rvert$ $R_2 = 0.05292$
$\text{model}_{1,1,2}$	$\lvert\ldots\ldots\rvert$	$\lvert\ldots\ldots\rvert$	$\lvert\ldots\ldots\rvert$
$\eta_1 = \lvert\ldots\ldots\rvert$	$F = \lvert\ldots\ldots\ldots\ldots\rvert$	$F = \lvert\ldots\ldots\ldots\ldots\rvert$	$F = \lvert\ldots\ldots\ldots\ldots\rvert$
$R_1 = 0.06401$	$Q_F = \lvert\ldots\ldots\ldots\ldots\rvert$	$Q_F = \lvert\ldots\ldots\ldots\ldots\rvert$	$Q_F = \lvert\ldots\ldots\ldots\ldots\rvert$
$\text{model}_{2,1,2}$	–	$\lvert\ldots\ldots\rvert$	$\lvert\ldots\ldots\rvert$
$\eta_1 = \lvert\ldots\ldots\rvert$	–	$F = \lvert\ldots\ldots\ldots\ldots\rvert$	$F = \lvert\ldots\ldots\ldots\ldots\rvert$
$R_1 = 0.05800$	–	$Q_F = \lvert\ldots\ldots\ldots\ldots\rvert$	$Q_F = \lvert\ldots\ldots\ldots\ldots\rvert$
$\text{model}_{3,1,2}$	–	–	$\lvert\ldots\ldots\rvert$
$\eta_1 = \lvert\ldots\ldots\rvert$	–	–	$F = \lvert\ldots\ldots\ldots\ldots\rvert$
$R_1 = 0.05434$	–	–	$Q_F = \lvert\ldots\ldots\ldots\ldots\rvert$

The sample size was $n = 514$.

Print this page. **Apply fisher.py** to the above results. In places marked $\lvert\ldots\ldots\rvert$ above, use a pen to **draw** missing arrows $\uparrow$ or $\leftarrow$ by hand, and **write** missing numbers $\lvert\ldots\ldots\rvert$ by hand using following accuracy

η_1 and η_2 (integer, e.g. $\eta_1 = 12$)

F (one decimal accuracy, e.g. $F = 14.9$)

Q_F (two first significant digits, e.g. $Q_F = 1.1 \times 10^{-15}$ or $Q_F = 0.0042$).

Scan your final hand written answers to pdf-file **FisherTable.pdf**.

Answer with less than 100 words, which model is the best model? This answer should be based on the two alternatives: "H_0 is rejected" or "H_0 is not rejected".

Send your **FisherTable.pdf** file and your **answer** to the assistant.