Practical Day 2

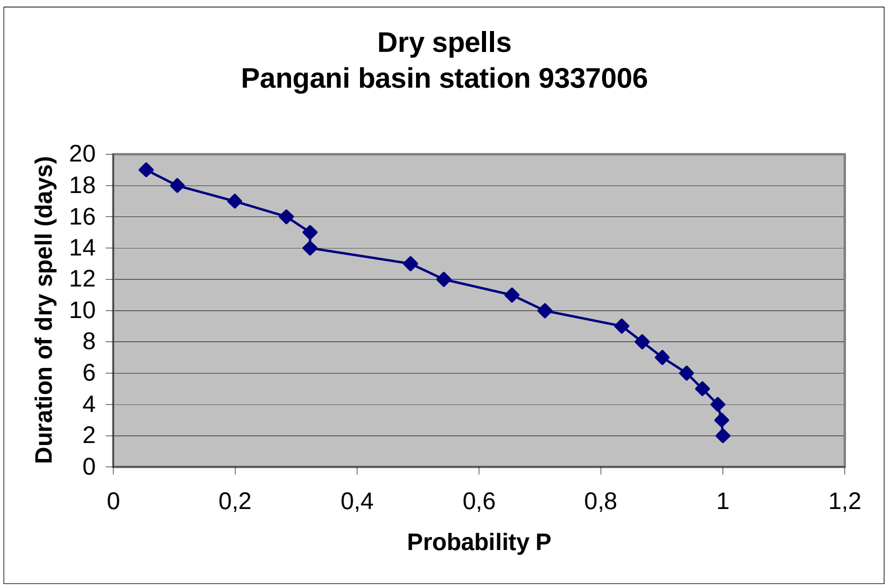
Exercise Dry Spell Analysis.

The occurrence of dry spells in the main wet season is analyzed through a frequency analysis of daily rainfall records of the rainfall station coded (9337006) in the Pangani Basin, Tanzania. The main wet season runs from 1st March to 31st May and consists of 92 days. Daily rainfall data available for 18 years is provided on blackboard.
Theory for this analysis is provided in the lecture notes chapter **1.2.4 Analysis of dry spells.**

1) Using an excel sheet define for all the years during the main wet season the number of dry spells of durations varying from 2 to 20 days.

2) Follow the procedure as described in the lecture note to come up with the probability that a dry spell longer than a certain duration occurs.

3) What is your conclusion on the possibilities for rainfed agriculture at this location.



spell duration t	number of spells i	accum spells I	start days per season n	total start days n*18 N	I/N p	1-p q	1-q^n P
2	58	166	91	1638	0,101343	0,898657	0,99994
3	24	108	90	1620	0,066667	0,933333	0,997989
4	24	84	89	1602	0,052434	0,947566	0,991716
5	10	60	88	1584	0,037879	0,962121	0,966563
6	9	50	87	1566	0,031928	0,968072	0,940578
7	5	41	86	1548	0,026486	0,973514	0,900588
8	4	36	85	1530	0,023529	0,976471	0,867862
9	10	32	84	1512	0,021164	0,978836	0,834181
10	3	22	83	1494	0,014726	0,985274	0,70809
11	5	19	82	1476	0,012873	0,987127	0,654378

12	2	14	81	1458	0,009602	0,990398	0,542298
13	5	12	80	1440	0,008333	0,991667	0,488015
14	0	7	79	1422	0,004923	0,995077	0,322841
15	1	7	78	1404	0,004986	0,995014	0,322849
16	2	6	77	1386	0,004329	0,995671	0,283987
17	2	4	76	1368	0,002924	0,997076	0,199523
18	1	2	75	1350	0,001481	0,998519	0,105234
19	1	1	74	1332	0,000751	0,999249	0,05406
20	0	0	73	1314			

Exercise Extreme Value Distribution and Mixed Distribution.

In this exercise an extreme value distribution will be fitted through the extreme values of runoff from the station Pavlovka Uborka in a river on the East Russian Coast. First of all it is assumed that the extreme values are homogeneous and originate from one and the same population. However, it is known that some of the extremes are the result of cyclones (named typhoons in this part of the world) that come to land from the Pacific Ocean. Therefore, secondly, a heterogeneous distribution will be assumed recognizing that the extremes originate from two different populations, i.e. extremes from cyclones and extremes from thunderstorms events. Hence the result is the combination of two distributions, the so called mixed distribution.

1) Compile a Gumbel type I extreme value distribution assuming that all the data belongs to one and the same distribution. Plot the data and the Gumbel distribution using excel, i.e. the reduced variate y (on the x-axis) against the extreme X (on the y-axis). For the procedure to create a Gumbel distribution see the lecture notes chapter **2.6 Flood Frequency analysis, Gumbel type I**.

Give values of the parameters for the equation:

$$X = X_m + s.\frac{(y - y_m)}{s_y} \text{ and } y = a(X - b)$$

2) Now recognizing that there are two populations compile for each of the populations the Gumbel type I distribution and plot the data and the two Gumbel distributions in the same graph you made for question 1).

Give for each population values of the parameters for the equation of the Gumbel distribution:

$$X_1 = X_{1,m} + s.\frac{(y_1 - y_{1,m})}{s_{1,y}} \text{ and } y_1 = a_1(X_1 - b_1)$$

$$X_2 = X_{2,m} + s.\frac{(y_2 - y_{2,m})}{s_{2,y}} \text{ and } y_2 = a_2(X_2 - b_2)$$

3) Compile the mixed distribution and plot the result in the same graph from the previous questions

Hint:

The mixed distribution reads as (see lecture notes **1.2.2 Mixed distribution**):

$$p(X > x) = p_1(X > x \mid C).p(C) + p_2(X > x \mid T).p(T)$$

For a Gumbel distribution it is known that:

$$p_1(X > x \mid C) = 1 - e^{-e^{-y_1}} \text{ and } p_2(X > x \mid T) = 1 - e^{-e^{-y_2}}$$

To create the curve of the mixed distribution every time select an extreme X, calculate the corresponding reduced variate y_1 and y_2 for each distribution and the corresponding $p_1(X>x,|C)$

and $p_2(X>x|T)$. This provides the basis to determine the probability $p(X>x)$ of the mixed distribution for the extreme X.

1)

$$y = a(X - b)$$

n=	52
y_m	0,5496
s_y	1,163
y_m/s_y	0,472571
X_m	291,7404
s	276,6377
a	0,004204
b	161,0095

$$\boxed{a = s_y / s}$$

$$\boxed{b = X_m - s.\frac{y_m}{s_y}}$$

2)

Thunderstorms

n=	34
y_m	0,5392
s_y	1,1236
y_m/s_y	0,479886
X_m	182,2794
s	102,5544
a	0,010956
B	133,065

$$y = a(X - b)$$

$$\boxed{a = s_y / s}$$

$$\boxed{b = X_m - s.\frac{y_m}{s_y}}$$

Cyclones

n=	18
y_m	0,5182
s_y	1,0396
y_m/s_y	0,498461
X_m	498,5
s	374,0875
a	0,002779
b	312,032