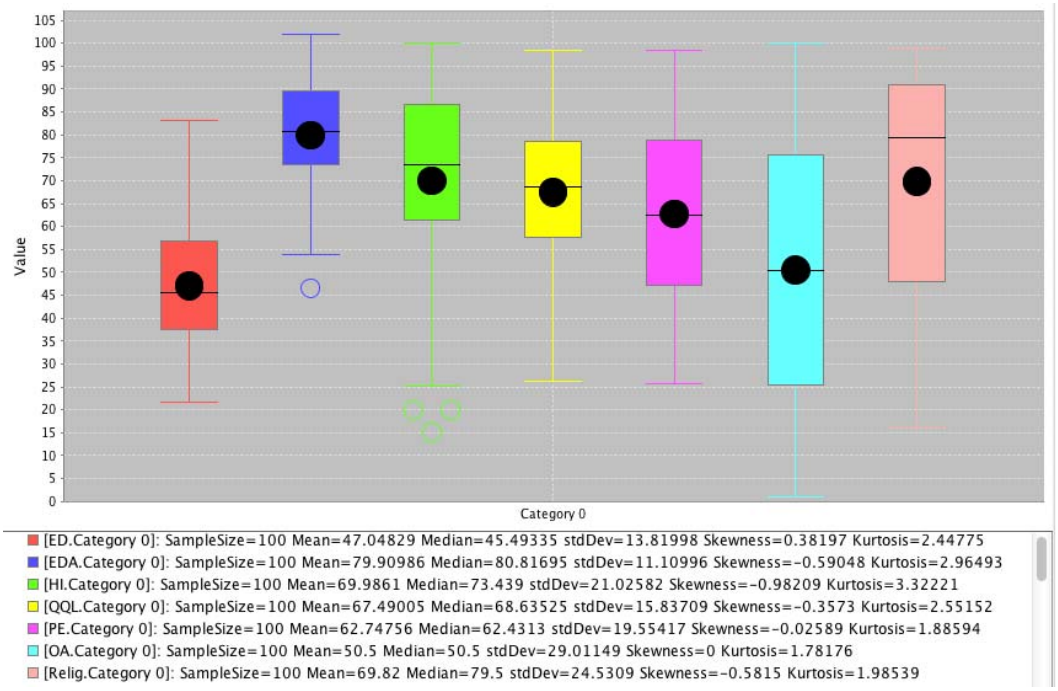
Scientific Methods for Health Sciences: Fundamentals (HS550): Fall 2014
http://www.socr.umich.edu/people/dinov/2014/Fall/HS550/
Homework 1[1] Solutions

Problem 1

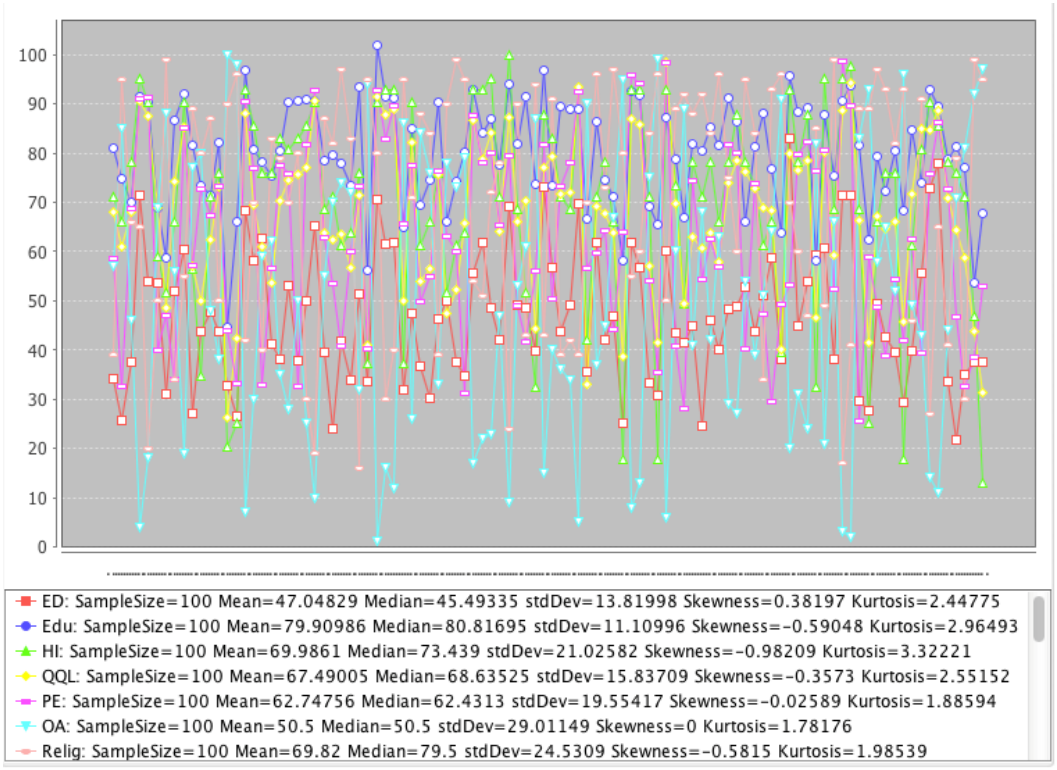
Boxplot



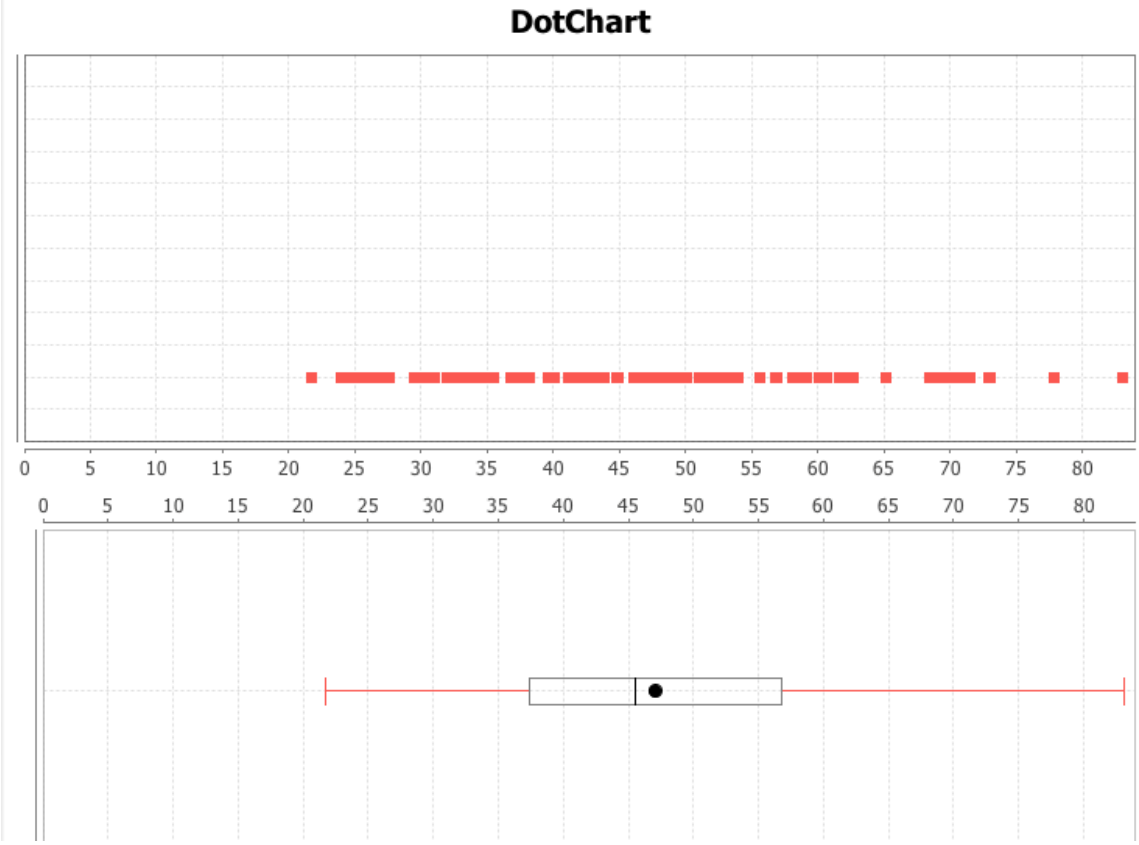
Scatterplot (ED vs. Edu)

[1] http://www.socr.umich.edu/people/dinov/2014/Fall/HS550/HWs.html

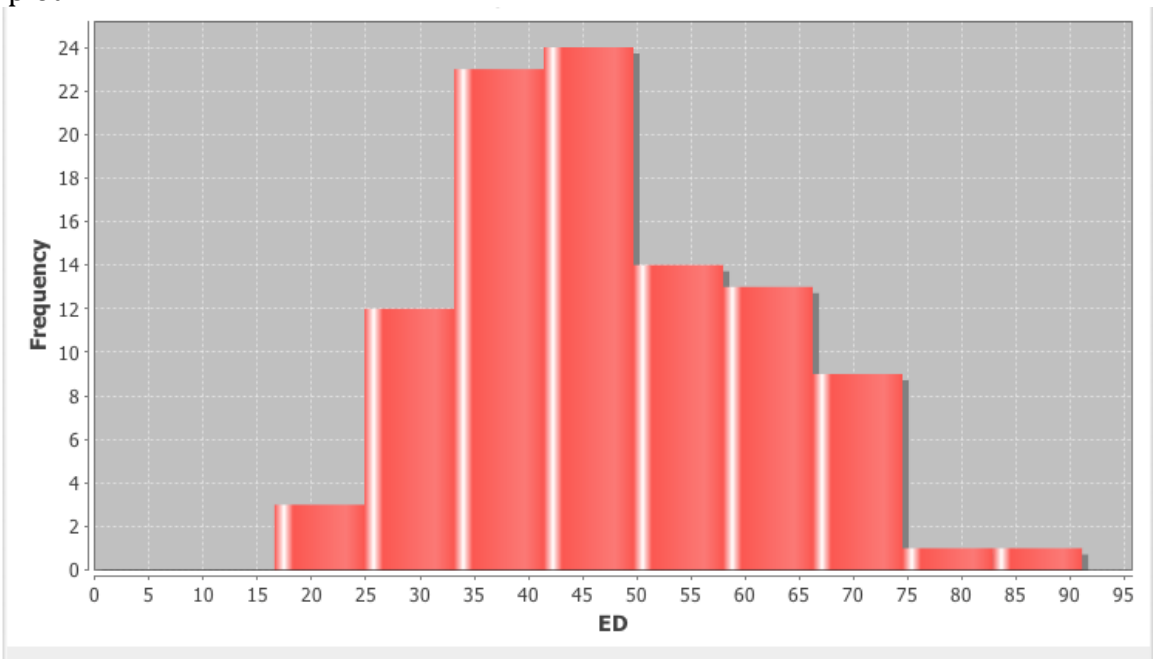
Line plot



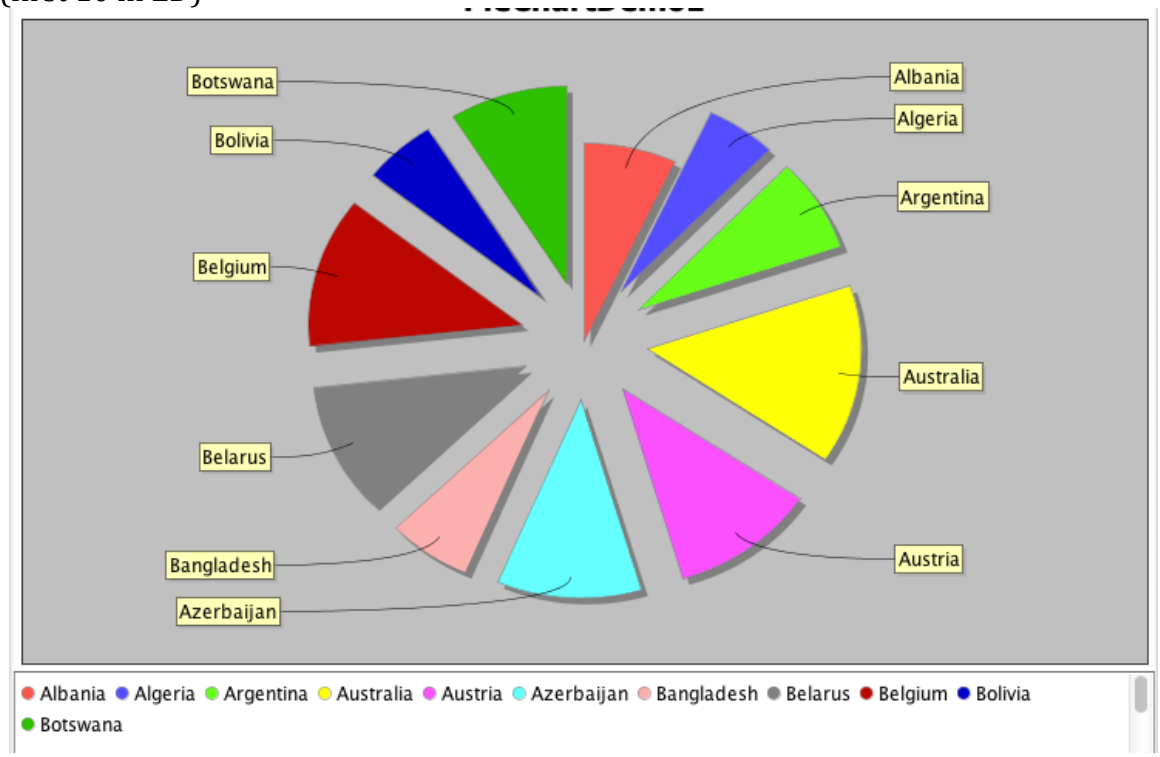
Dot plot: (ED)



Histogram plot:



Pie chart: (first 10 in ED)



Problem 2

In R:

u <- runif(100,-1,1)

-0.50593920 -0.05379413 0.44176964 -0.99069946 0.40658220 -0.31252493 -0.57072538 0.44484457 -0.40816991 -0.88165310 0.09429955 0.34787585 -0.47047205 0.29800894 0.64148757 0.13778942 0.91318412 0.39542682 -0.60454238 -0.84686276 -0.40352041 -0.36349603 -0.09941884 0.64971226 0.32414321 -0.52847884 0.87661907 -0.62739848 -0.52663591 -0.17378271 -0.18881219 -0.43301409 -0.15303192 -0.28445506 -0.63678817 -0.57669355 -0.13199231 0.86270131 -0.49473818 0.56465117 -0.10473786 0.64582830 0.44609494 -0.60037943 0.92062335 -0.86459057 -0.63427335 0.75332309 -0.56604354 -0.42338155 -0.81942114 -0.39320780 0.20493439 0.78483884 -0.78024256 -0.79813815 0.88229600 -0.16786827 0.18190380 -0.61330601 -0.03383999 0.85888243 -0.19559880

0.02812067 -0.84928048 0.98407000 -0.86266605 0.23725486 -0.20089859 0.64472711 0.08329493 0.16689940 0.46343894 -0.77247763 0.68432733 0.74841489 -0.60365680 -0.65197992 -0.48007512 0.23916335 -0.72705192 0.81567585 -0.67305070 0.08844869 0.66838513 0.08198096 0.64961482 -0.39512221 -0.45261468 -0.94425210 -0.65270209 0.19361020 0.74267481 -0.94175317 -0.11434287 -0.01814690 -0.34746315 -0.17098867 -0.68688800 -0.63284361

n <- rnorm(100,0,1)

-0.096100238 1.260403830 -0.230104571 0.357100277 -0.079650421 0.868287425 -0.377432404 -0.821750114 -0.908713377 0.102564483 -0.288183707 1.334516118 -1.069111249 -0.113446509 1.518886876 1.609114036 -0.714385454 0.647725640 1.084468428 -0.551301712 -1.529763370 -1.291841116 -0.909384921 -1.357114470 2.399599664 0.559480837 0.487697500 -1.267566270 -0.355004969 0.682719220 -0.234782637 0.504640148 0.081468802 -0.305563004 -1.228383306 0.021333769 -0.688850226 -1.877900929 1.349589573 1.654438659 -0.397858066 0.867576726 1.237258281 -0.688246993 -1.143944018 0.357583515 -1.489252417 -1.252676722 1.997605258 1.008110174 1.568881345 -0.791919409 1.059703488 -1.018645617 1.205648659 1.047044818 0.256814671 1.638629720 0.570031167 0.141554260 0.643774910 1.433482245 1.753241883 1.271050050 -0.905388022 -0.330329565 -0.551817195 0.435087151 -0.006234462 -0.652073400 0.670984932 -1.686999371 1.026125904 -1.991309625 -0.042824035 -0.723156408 -1.154745566 0.642544225 0.039574667 -0.068064154 -2.101759155 0.427014285 1.263001591 -1.584336156 -0.558768131 0.245024422 0.247409991 0.117726718 -0.842044920 0.534978243 -0.353848080 0.950652558 0.841082783 -0.348300151 -0.600919570 -0.380803807 -0.827720846 -0.892678011 0.581127557 0.304570636

e <- rexp(100,1)

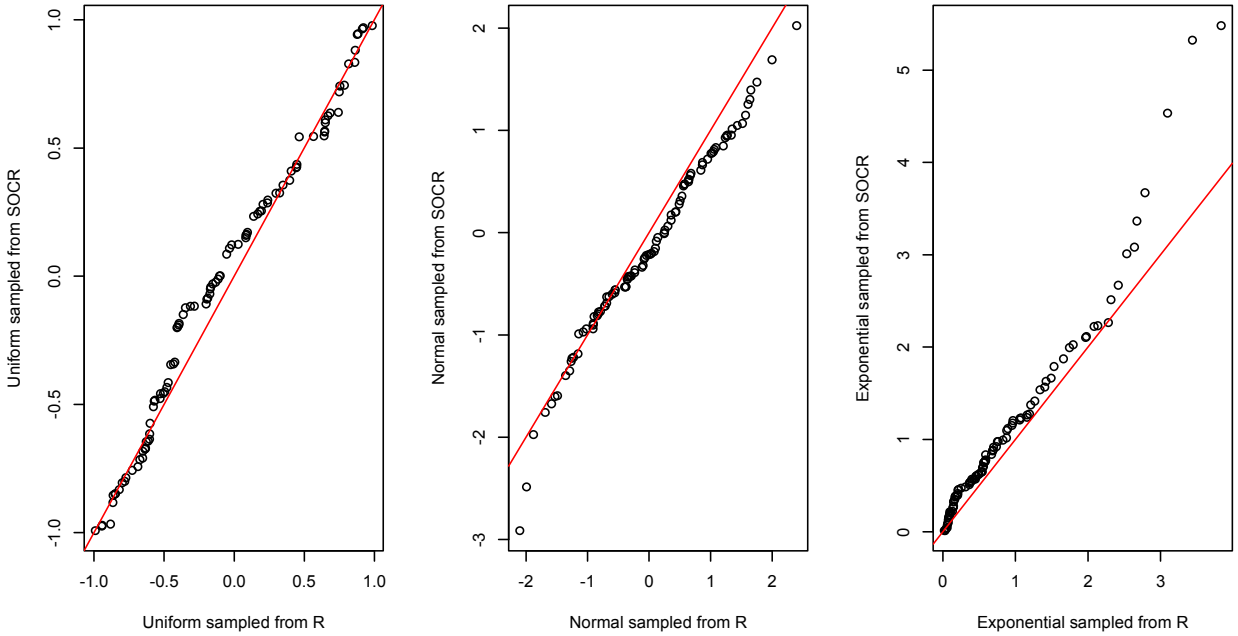
0.09463080 2.28219116 0.46285911 1.05897714 0.06090467 1.07935067 0.88051893 0.69744207 0.16772525 0.31299042 1.74572473 3.83682764 0.42885591 0.58893768 2.64008244 1.66295398 0.49656974 1.79640547 0.16453788 0.21628215 0.67045882 0.19635792 0.07399550 0.45460535 0.57545902 2.41817958 0.87859975 0.20374237 0.76272076 0.37563896 3.43941399 0.06500675 0.01862754 0.95181110 0.69019103 1.16242009 0.08899573 0.17288592 1.53227979 0.36694587 0.89541833 1.42144981 0.07399724 0.14453579 0.20798946 0.03205865 0.13257516 2.31860297 0.10900662 0.05972697 0.44971323 0.04441783 0.55784237 0.09547163 0.03832134 1.33820923 2.08370463 0.14077887 0.14844905 0.52657661 2.67408017 0.56300917 1.21160227 1.19428141 0.20690717 0.25204410 1.06510496 0.58917251 0.53049698 2.13436372 0.07825004 0.39749462 0.36415085 1.96743773 0.96613578 0.14856824 2.78508491 0.96300925 0.73858739 1.15904640 3.09794967 0.54741728 2.53454679 0.15047202 1.40405396 0.58683101 0.39937198 0.08649058 0.68655517 1.97795280 0.82860098 1.26629860 0.54225694 0.48385269 0.75053439 0.68520377 0.56688020 1.49221988 0.09558541 0.08101499

Data sampled organized in table:

Uniform(-1,1)		Normal(0,1)		Exponential(1)	
R	**SOCR**	**R**	**SOCR**	**R**	**SOCR**
-0.5059392	-0.084879261	-0.096100238	-0.426537882	0.0946308	0.012702814
-0.05379413	-0.854872108	1.26040383	0.579670166	2.28219116	0.263912412
0.44176964	-0.477034567	-0.230104571	0.200371322	0.46285911	1.663827128
-0.99069946	0.28011554	0.357100277	-0.903178323	1.05897714	0.922350048
0.4065822	-0.715573542	-0.079650421	-1.184241623	0.06090467	0.885376245
-0.31252493	-0.805304722	0.868287425	-1.75714574	1.07935067	0.351662356
-0.57072538	0.545175384	-0.377432404	1.302655984	0.88051893	1.203670013
0.44484457	-0.049676652	-0.821750114	0.4572448	0.69744207	0.834207322
-0.40816991	0.609453874	-0.908713377	-0.530627713	0.16772525	0.993175473
-0.8816531	0.108588004	0.102564483	-0.794554931	0.31299042	1.150873965
0.09429955	-0.848402972	-0.288183707	1.013814629	1.74572473	0.916589722
0.34787585	0.827957607	1.334516118	-0.423531849	3.83682764	0.624140716
-0.47047205	-0.344571234	-1.069111249	1.047806175	0.42885591	1.017157148
0.29800894	-0.574054258	-0.113446509	0.560007308	0.58893768	3.082669224
0.64148757	-0.123338896	1.518886876	-0.811446	2.64008244	0.752658142
0.13778942	0.435932712	1.609114036	-1.593926322	1.66295398	1.23304198
0.91318412	0.25324889	-0.714385454	-0.528827529	0.49656974	0.025816568
0.39542682	-0.849278494	0.64772564	-0.559597605	1.79640547	0.747529155

Uniform(-1,1)		Normal(0,1)		Exponential(1)	
-0.60454238	0.149837886	1.084468428	-1.258623612	0.16453788	0.181251525
-0.84686276	-0.785690072	-0.551301712	0.064519006	0.21628215	1.229104675
-0.40352041	0.285711411	-1.52976337	0.848203022	0.67045882	0.651312188
-0.36349603	0.423160657	-1.291841116	-0.600873541	0.19635792	1.415649672
-0.09941884	0.162931121	-0.909384921	0.518185287	0.0739955	0.39289295
0.64971226	0.233525648	-1.35711447	-0.221361098	0.45460535	1.210144106
0.32414321	-0.184106572	2.399599664	0.928533631	0.57545902	5.481635557
-0.52847884	-0.971544359	0.559480837	-0.688294562	2.41817958	0.695663376
0.87661907	-0.190268984	0.4876975	0.025260487	0.87859975	0.880729899
-0.62739848	0.744581569	-1.26756627	-1.973644778	0.20374237	0.617972072
-0.52663591	0.718585738	-0.355004969	-0.938322377	0.76272076	0.012361678
-0.17378271	0.543418373	0.68271922	-0.083774732	0.37563896	0.752260093
-0.18881219	0.94271617	-0.234782637	-0.047918984	3.43941399	5.323511093
-0.43301409	-0.335109915	0.504640148	1.256146409	0.06500675	0.221524293
-0.15303192	0.833884756	0.081468802	-1.227733523	0.01862754	1.537131544
-0.28445506	-0.96665951	-0.305563004	-0.32125817	0.9518111	1.238668082
-0.63678817	0.88094772	-1.228383306	0.774145459	0.69019103	0.154773244
-0.57669355	-0.508952265	0.021333769	1.148035645	1.16242009	2.025242783
-0.13199231	-0.117251085	-0.688850226	0.357282838	0.08899573	0.643640868
0.86270131	-0.614959566	-1.877900929	-0.337229905	0.17288592	0.215589052
-0.49473818	0.355251785	1.349589573	1.471512505	1.53227979	0.386295228
0.56465117	-0.415114247	1.654438659	0.663208362	0.36694587	0.071712954
-0.10473786	-0.832469753	-0.397858066	0.832538281	0.89541833	0.185755331
0.6458283	0.598423611	0.867576726	9.25E-05	1.42144981	2.231541183
0.44609494	-0.108760688	1.237258281	-0.98832638	0.07399724	2.513484628
-0.60037943	0.296727207	-0.688246993	0.718497986	0.14453579	0.184179102
0.92062335	0.426194291	-1.143944018	-0.184411041	0.20798946	0.56975754
-0.86459057	-0.342289321	0.357583515	-1.672474843	0.03205865	0.646745753
-0.63427335	-0.992333768	-1.489252417	-0.427874493	0.13257516	1.265916158
0.75332309	-0.030144496	-1.252676722	1.69001302	2.31860297	1.993571126
-0.56604354	-0.882854976	1.997605258	-0.77646063	0.10900662	1.790368502
-0.42338155	0.964816834	1.008110174	0.947004722	0.05972697	1.276464114
-0.81942114	0.547089084	1.568881345	1.067705749	0.44971323	0.214572746
-0.3932078	-0.487219511	-0.791919409	-0.21034179	0.04441783	0.451915706
0.20493439	-0.117725623	1.059703488	-1.34970009	0.55784237	0.323853686
0.78483884	0.741192885	-1.018645617	-0.586316487	0.09547163	0.704554747
-0.78024256	-0.042275248	1.205648659	0.953384638	0.03832134	0.77818997
-0.79813815	-0.069036715	1.047044818	-0.455638049	1.33820923	0.765225238
0.882296	0.969059038	0.256814671	0.173263574	2.08370463	3.67085982
-0.16786827	-0.709232134	1.63862972	-0.391250157	0.14077887	3.365004245
0.1819038	0.409876877	0.570031167	-1.396222523	0.14844905	1.117474658
-0.61330601	0.977241211	0.14155426	0.471251305	0.52657661	2.114143395

Uniform(-1,1)		Normal(0,1)		Exponential(1)	
-0.03383999	-0.011769719	0.64377491	-0.152490724	2.67408017	0.266409332
0.85888243	0.56545964	1.433482245	-0.624860471	0.56300917	0.37606896
-0.1955988	-0.742757446	1.753241883	0.814177996	1.21160227	0.043038267
0.02812067	-0.089728215	1.27105005	-0.201023113	1.19428141	3.010681681
-0.84928048	-0.458253654	-0.905388022	0.121959205	0.20690717	0.415647688
0.98407	0.086049629	-0.330329565	0.793579956	0.2520441	0.978720655
-0.86266605	-0.457131092	-0.551817195	-0.24297412	1.06510496	1.566515016
0.23725486	0.242892434	0.435087151	-0.770315471	0.58917251	2.104400225
-0.20089859	-0.023501137	-0.006234462	0.311908555	0.53049698	0.328343094
0.64472711	0.945044885	-0.6520734	0.205403167	2.13436372	0.571062112
0.08329493	0.001479472	0.670984932	-0.214113241	0.07825004	0.487063926
0.1668994	-0.682628812	-1.686999371	0.610471389	0.39749462	4.533586859
0.46343894	-0.197611054	1.026125904	0.779818409	0.36415085	0.129522443
-0.77247763	-0.645520887	-1.991309625	-0.36514444	1.96743773	0.457821995
0.68432733	0.32434535	-0.042824035	-0.718153989	0.96613578	0.87593022
0.74841489	0.562101511	-0.723156408	-0.626540245	0.14856824	0.150780211
-0.6036568	0.15911127	-1.154745566	0.521255285	2.78508491	0.606415485
-0.65197992	4.33E-04	0.642544225	-2.484558893	0.96300925	0.099190473
-0.48007512	-0.670171374	0.039574667	-0.719251304	0.73858739	0.396596264
0.23916335	-0.635627242	-0.068064154	-1.601727466	1.1590464	0.278139495
-0.72705192	0.124187576	-2.101759155	-0.452928031	3.09794967	0.529866509
0.81567585	-0.672621282	0.427014285	-0.819756429	0.54741728	2.223723059
-0.6730507	0.122178053	1.263001591	0.498547327	2.53454679	2.265495747
0.08844869	0.636425341	-1.584336156	-0.528148507	0.15047202	1.096184736
0.66838513	0.625494343	-0.558768131	2.024655122	1.40405396	0.572350493
0.08198096	-0.200706282	0.245024422	-2.912687182	0.58683101	1.176484056
0.64961482	0.323111374	0.247409991	0.94579322	0.39937198	0.601165838
-0.39512221	-0.64292111	0.117726718	-0.00979657	0.08649058	0.976324828
-0.45261468	0.254732173	-0.84204492	1.394610175	0.68655517	0.170630317
-0.9442521	0.170792621	0.534978243	-0.559080137	1.9779528	0.035984851
-0.65270209	-0.974829242	-0.35384808	-0.886508363	0.82860098	2.670271205
0.1936102	0.638798383	0.950652558	0.953449405	1.2662986	0.538763532
0.74267481	0.373988071	0.841082783	-0.96970564	0.54225694	0.549017135
-0.94175317	-0.149614912	-0.348300151	-0.257148879	0.48385269	0.83797312
-0.11434287	-0.757208843	-0.60091957	-0.939877491	0.75053439	0.09144773
-0.0181469	-0.799857446	-0.380803807	0.68606761	0.68520377	0.511205309
-0.34746315	-0.434255653	-0.827720846	-1.218069158	0.5668802	1.629748565
-0.17098867	-0.450729562	-0.892678011	0.278541248	1.49221988	1.374295172
-0.686888	-0.673525013	0.581127557	0.477363253	0.09558541	0.47326565
-0.63284361	-0.483883509	0.304570636	-0.434153991	0.08101499	1.873111133



For the uniform distribution, the curve fit the red line pretty well except for some small wriggle about the line so we can say that the two samples sampled from R and SOCR are similar to each other. This is also the case with normal distribution where there is a slight heavy tail at the ends, which is reasonable since the normal distribution follow a bell shape with the majority of the points falling around the origin and outliers may exist and the frequency would largely depend on the standard deviation. However, for the exponential distribution, the curve seem to have a very heavy tail at the right end which matches the expectation since exponential function with parameter 1 would generate samples whose value can be pretty big instead of concentrating within the range of a certain boundary like uniform and normal distributions do, both of which are symmetric distributions.

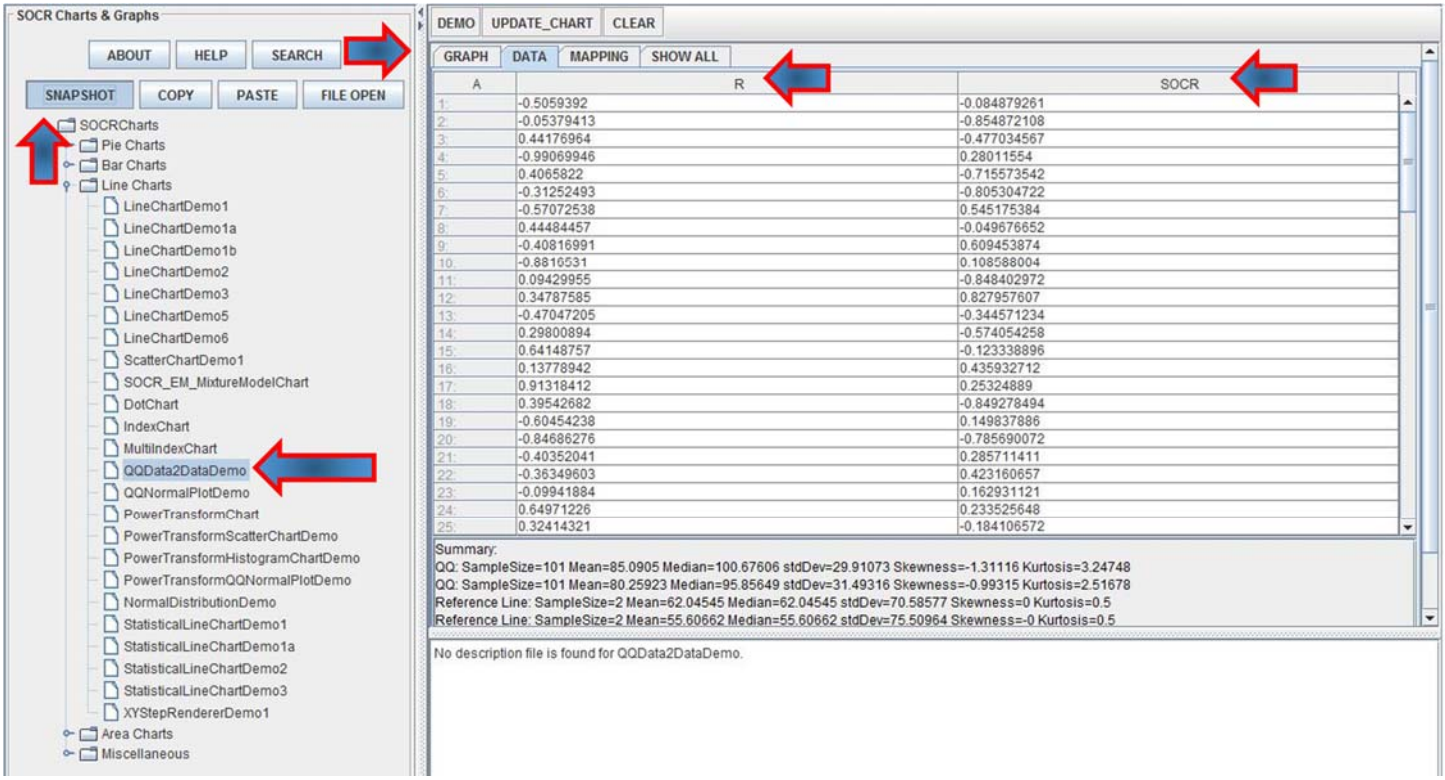
Anyway, we don't expect random list of 100 samples sampled from different tools to match perfectly well given the existence of chances and randomness. Though we do expect the patterns (in the QQ Plot the quantiles) to be correlated, according to the characteristics of the specific distribution.

In R:

```
data <- read.csv('Problem2.csv',header=T)
attach(data)
colnames(data)=c('U_R','U_S','N_R','N_S','E_R','E_S')
par(mfrow=c(1,3))
qqplot(U_R,U_S,xlab='Uniform sampled from R',ylab='Uniform sampled from SOCR')
abline(0,1,col='red')
qqplot(N_R,N_S,xlab='Normal sampled from R',ylab='Normal sampled from SOCR')
abline(0,1,col='red')
qqplot(E_R,E_S,xlab='Exponential sampled from R',ylab='Exponential sampled from SOCR')
abline(0,1,col='red')
```

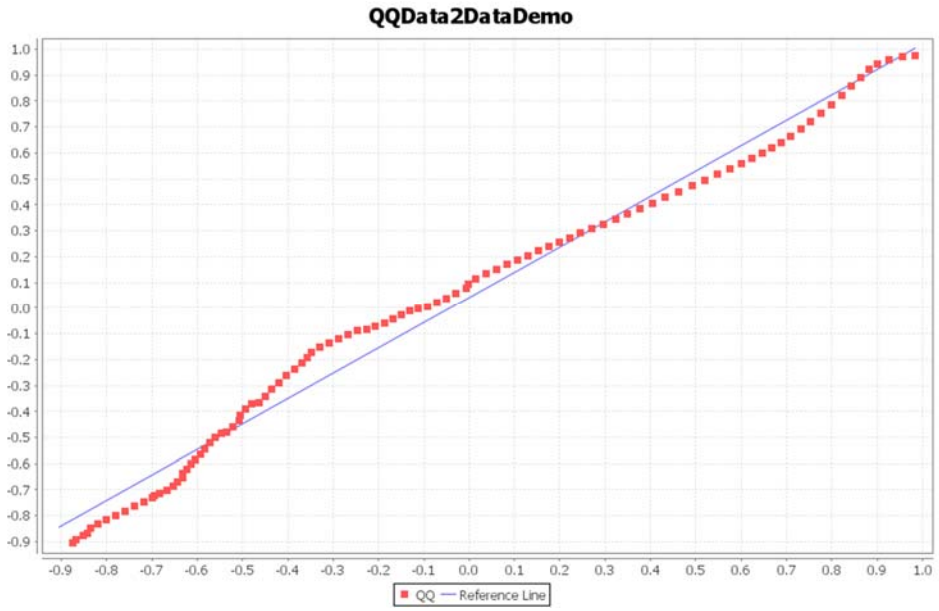
Using SOCR QQ Data to Data Plot (http://www.socr.umich.edu/html/cha/SOCR_Charts.html)

Copy-paste pairs of data (randomly generated samples from R and SOCR) for each of the 3 distributions.

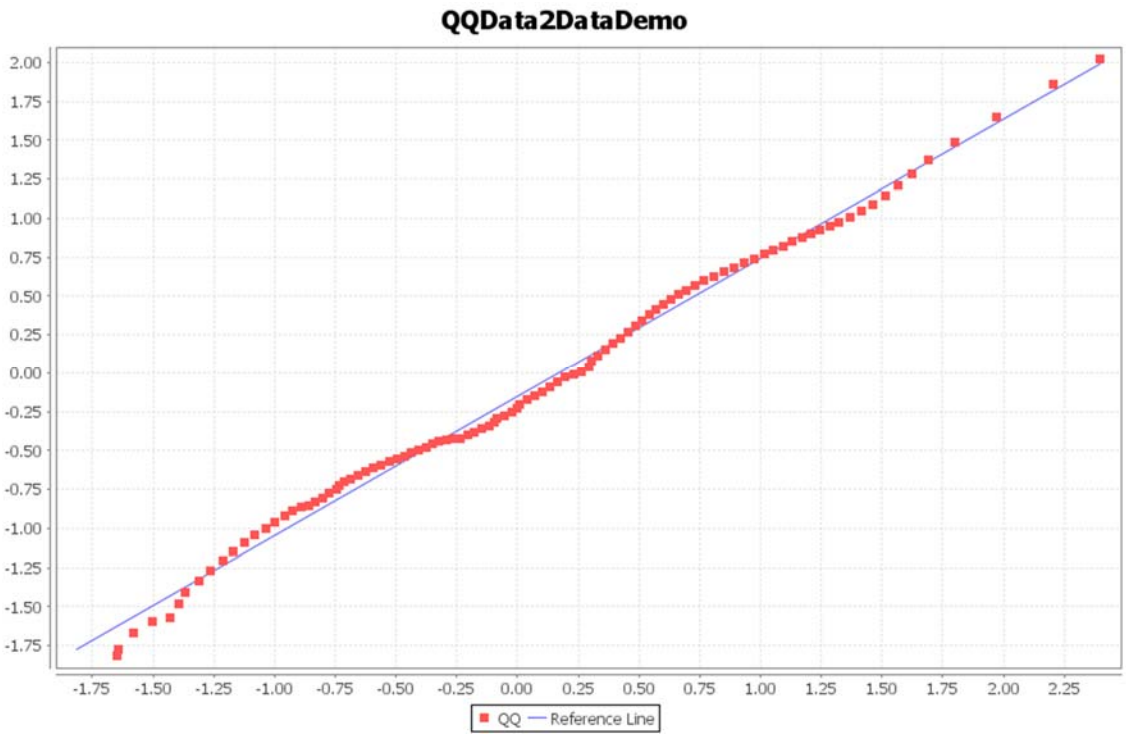


Then click the mapping tab and map the R and SOCR column headings into the 2 bins (X and Y axes). Finally click "Update Chart" to plot the QQ plot (quartiles of R and SOCR data are on the X and Y axes, respectively). We can change the Chart Titles and label the axes (using right-click and setting the chart properties), however we did not do that here to avoid possible confusion.

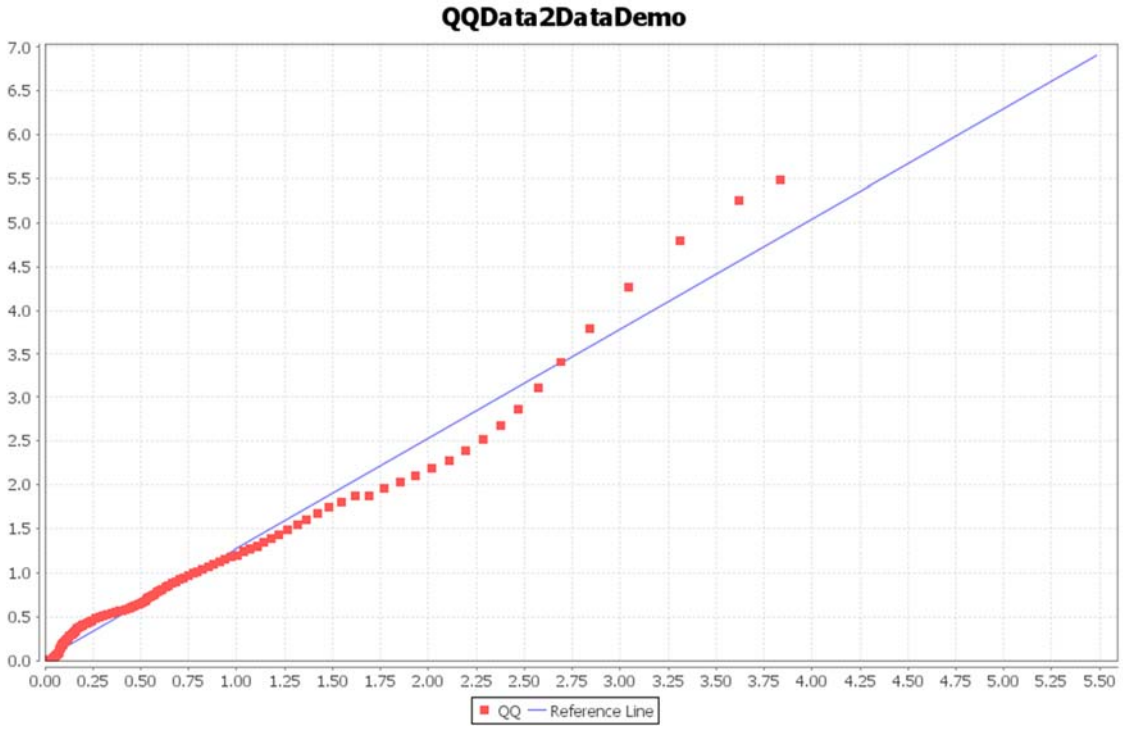
Uniform(0,1)



Normal(0,1)



Exponential(1)



Problem 3

Type	Site			Totals
	Head and Neck	Trunk	Extremities	
Hutchinson's melanomic freckle	22	2	10	34
Superficial	16	54	115	185
Nodular	19	33	73	125
Indeterminant	11	17	28	56
Column Totals	68	106	226	400

- P(Extremeities | Nodular) = 73/125=0.584.
- P(superficial | Trunk)=54/106=0.509434.

Problem 4

		Heart Attack (HA)		Total
		Yes	**No**	
Smoking (S)	**Yes**	33	18	51
	No	167	182	349
Total		200	200	400

Odds Ratio of heart attach relative to smoking:
OR=(33*182)/(18*167)=1.99804. The odds of heart attach is about 2 times greater for smoking persons compared to non-smoking persons in the study.
SE(ln(OR))=\sqrt{1/n_{1,1}+1/n_{1,2}+1/n_{2,1}+1/n_{2,2}}=
\sqrt{1/33+1/18+1/167+1/182}=0.3119954.
The 95% CI of log(OR) is
ln(OR) \neq z_{\alpha /2}SE(ln(OR))=ln(1.99804) \neq 1.96*0.33119954 = (0.0430156, 1.341318).

in R:

```
> log(1.99804,base=exp(1))+1.96*0.3119954
[1] 1.303678
> log(1.99804,base=exp(1))-1.96*0.3119954
[1] 0.08065572

> # these are the log limits, we need to exponentiate to get the standard unit limits
> exp(1.303678)
[1] 3.682816
> exp(0.08065572)
[1] 1.083998
```

So, the 95% CI for OR: (1.083998, 3.682816).