

Pensieve header: Data analysis for Maya Bar-Hillel.

In the file attached there are the results of 280 respondents, each of whom rated 28 variables on a 1-to-5 scale.

I'd like a Monte Carlo to assess the correlation between the means given each variable by two random halves of the 280 people.

So: divide the 280 into 2 groups of 140 each, at random.

Compute the mean rating for each of the 28 variables, separately for each half.

Compute Pearson's r between these mean ratings.

Repeat N times (one hundred?).

Calculate the mean r.

I suspect it will be close to 1, over and over and over again.

Initialization

```
In[ ]:= SetDirectory["C:\\drorbn\\AcademicPensieve\\People\\BarHillel"]
```

```
Out[ ]:= C:\drorbn\AcademicPensieve\People\BarHillel
```

```
In[ ]:= tab = Import["Monte.xlsx"][[1];
Dimensions[tab]
```

```
Out[ ]:= {280, 28}
```

```
In[ ]:= Av[L_List] := Mean[Select[L, NumberQ]];
```

Comparing the first 140 with the last 140

```
In[ ]:= {A = Range[140], B = Complement[Range[280], A]};
{AvA = Av /@ Transpose@tab[[A]], AvB = Av /@ Transpose@tab[[B]]}
basecor = Correlation[AvA, AvB]
```

```
Out[ ]:= {{2.56589, 1.56349, 3.13973, 3.66423, 2.36296, 2.56391, 2.63846, 2.56489, 2.72308, 3.12977,
3.42366, 3.59848, 3.16667, 4.24627, 2.88372, 3.32061, 3.29924, 2.632, 2.8626,
3.09848, 3.39394, 3.35606, 3.23438, 2.19084, 2.17424, 2.66412, 2.58209, 2.87879},
{2.37209, 1.59375, 2.53846, 3.17557, 2.18321, 2.43846, 2.74242, 2.31496, 2.47619,
2.73282, 2.99225, 3.14615, 2.63636, 4.21805, 2.70313, 3.25984, 3.33333, 2.55814,
2.86364, 3.09848, 3.22137, 3.21053, 3.312, 2.44186, 2.37209, 2.87023, 2.7803, 2.83077}}
```

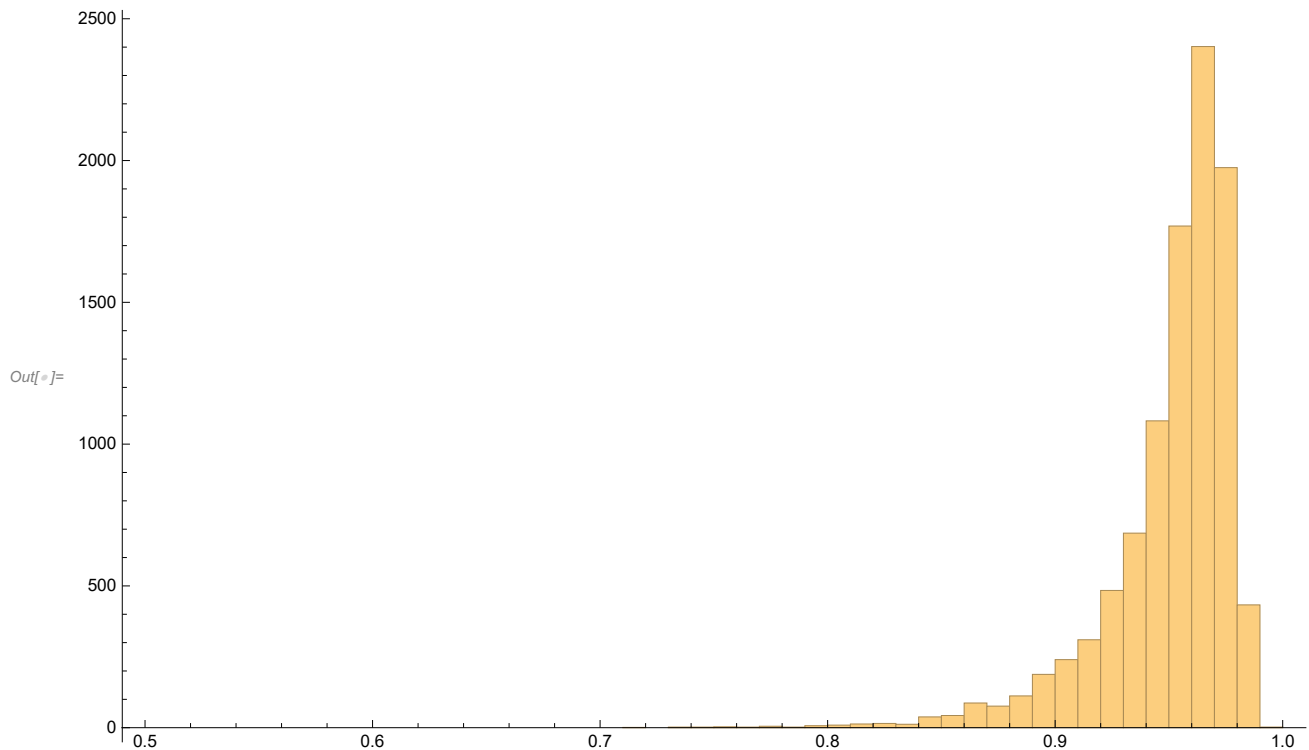
```
Out[ ]:= 0.901919
```

Monte-Carlo within 1..280, 10,000 runs:

```
In[ ]:= Short[cors = Sort@Table[
  {A = Sort@RandomSample[Range[280], 140], B = Complement[Range[280], A]};
  AvA = Av /@ Transpose@tab[A];
  AvB = Av /@ Transpose@tab[B];
  Correlation[AvA, AvB],
  {10 000}
]]
```

```
Out[ ]:= Short= {0.716528, 0.731318, 0.733987, 0.742962, <<9993>>, 0.989979, 0.990871, 0.991736}
```

```
In[ ]:= Histogram[cors, {0.5, 1, 0.01}, ImageSize -> 640]
```



```
In[ ]:= Mean[cors]
```

```
Out[ ]:= 0.95098
```

```
In[ ]:= Count[cors, _? (# < basecor &)] / Length[cors] // N
```

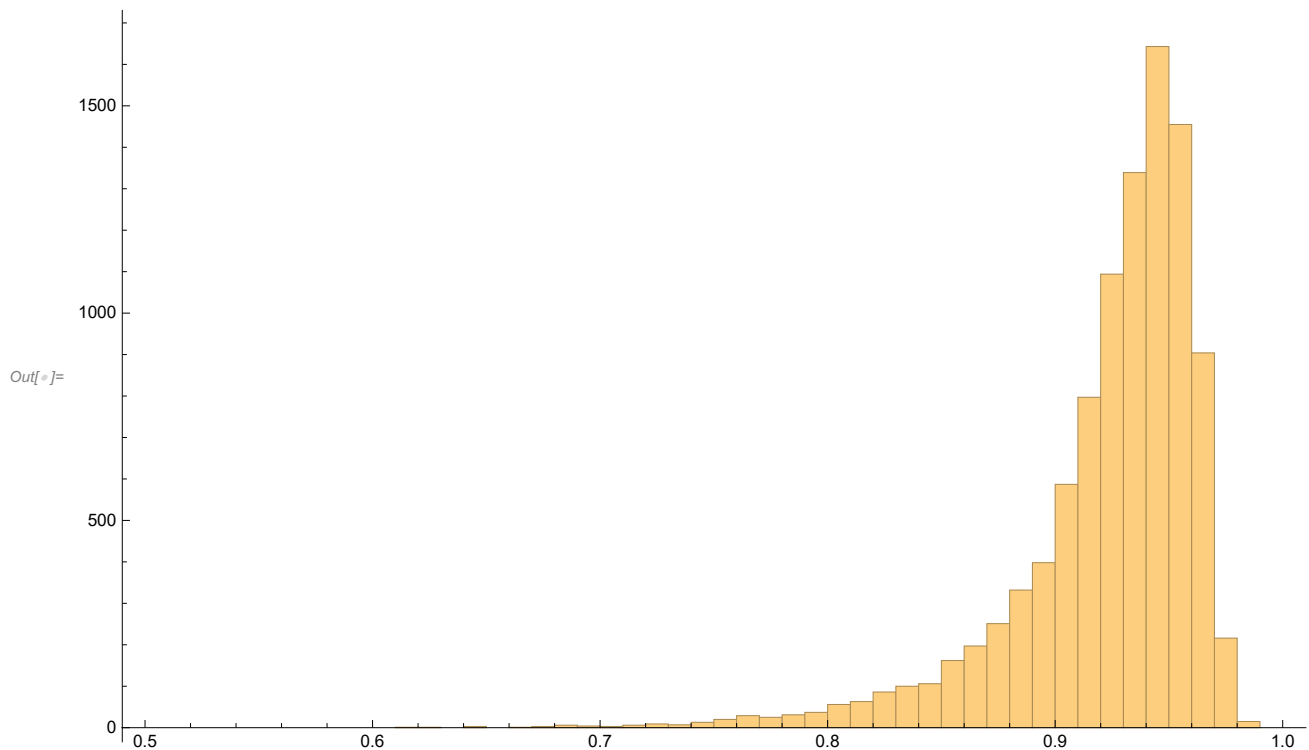
```
Out[ ]:= 0.0655
```

Monte-Carlo within 1..140, 10,000 runs:

```
In[ ]:= Short[cors = Sort@Table[
  {A = Sort@RandomSample[Range[140], 70], B = Complement[Range[140], A]};
  AvA = Av /@ Transpose@tab[A];
  AvB = Av /@ Transpose@tab[B];
  Correlation[AvA, AvB],
  {10 000}
]]
```

```
Out[ ]:= Short= {0.613365, 0.624736, 0.642722, 0.644107, <<9993>>, 0.984205, 0.984444, 0.985826}
```

```
In[ ]:= Histogram[cors, {0.5, 1, 0.01}, ImageSize -> 640]
```



```
In[ ]:= Mean[cors]
```

```
Out[ ]:= 0.923445
```

```
In[ ]:= Count[cors, _? (# < basecor &)] / Length[cors] // N
```

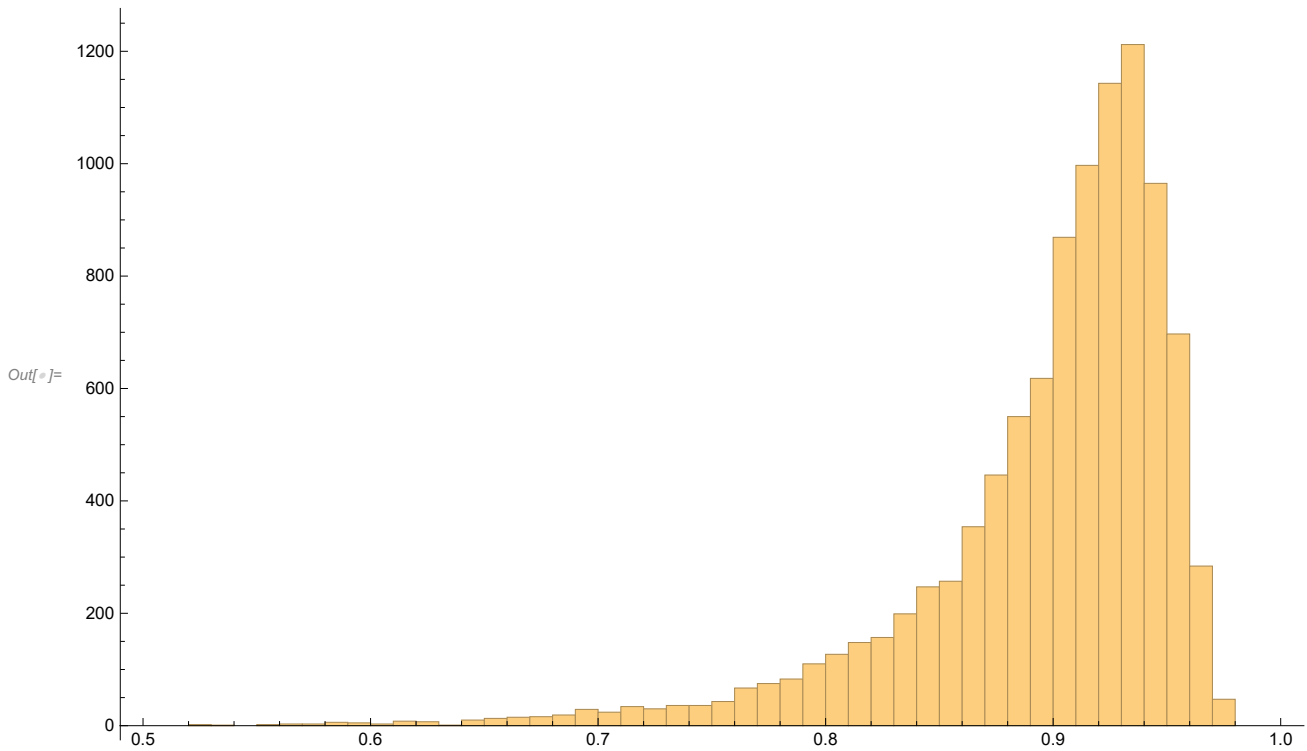
```
Out[ ]:= 0.2055
```

Monte-Carlo within 141..280, 10,000 runs:

```
In[ ]:= Short[cors = Sort@Table[
  {A = Sort@RandomSample[Range[141, 280], 70], B = Complement[Range[141, 280], A]};
  AvA = Av /@ Transpose@tab[A];
  AvB = Av /@ Transpose@tab[B];
  Correlation[AvA, AvB],
  {10 000}
]]
```

```
Out[ ]:= Short= {0.40477, 0.471789, 0.520285, 0.520806, <<9993>>, 0.976575, 0.976827, 0.977084}
```

```
In[ ]:= Histogram[cors, {0.5, 1, 0.01}, ImageSize -> 640]
```



```
In[ ]:= Mean[cors]
```

```
Out[ ]:= 0.897475
```

```
In[ ]:= Count[cors, _? (# < basecor &)] / Length[cors] // N
```

```
Out[ ]:= 0.3955
```