Statistical Doubles – a Quiz

Gaj Vidmar

Countless things in nature, society and universe exist in pairs. Likewise, many sports can be played in doubles. One of those is tennis, and its oldest and most prestigious tournament has been held in Wimbledon for nearly a century and a half, which is the inspiration for this writing.

But it is not about sports statistics. It is about gentlemen's doubles, but we will neither attempt to predict this year's Wimbledon champions nor compare Mc Millan-Hewitt, Newcombe-Roche, McEnroe-Flemming, the Woodies and other legends (incidentally, my favourites have always been Fitzgerald-Järryd). I merely present examples of doubles as co-authors of, or name-givers to, statistical methods. Below is a list of 32 (to stay with the analogy of tennis). On the right-hand side, brief descriptions are listed in scrambled order, so the readers are invited to test their knowledge by **connecting the authors with the methods**:

1	Jensen–Shannon	test for stratified 2×2 tables	1
2	Kullback–Leibler	test for outlier(s)	2
3	Shapiro–Wilk	agreement plot	3
4	Anderson–Darling	estimator for kernel regression	4
5	Cramér–von Mises	non-parametric two-sample test of scale	5
6	Jarque–Bera	monotonic power-transform of data	6
7	Kolmogorov–Smirnov	less known non-parametric two-sample test of scale	7
8	Mann–Whitney	two-independent-samples test of medians	8
9	Fligner–Policello	non-parametric estimator of cumulative hazard	9
10	Ansari–Bradley	non-parametric regression for method comparison	10
11	Siegel–Tukey	test of a probability distribution	11
12	Kaplan–Meier	statistic to detect autocorrelation of regression residuals	12
13	Nelson–Aalen	test of normal distribution	13
14	Fleming–Harrington	root-finding numerical algorithm	14
15	Tietjen–Moore	goodness-of-fit test for logistic regression	15
16	Box–Cox	variant of logrank two-sample test of survival	16
17	Bland–Altman	less used estimator of survival function	17
18	Passing–Bablok	two-sample test of survival	18
19	Tarone–Ware	test for heteroskedasticity in a linear regression	19
20	Peto–Peto	goodness-of-fit criterion for cumulative distribution	20
21	Mantel–Haenszel	symmetric relative entropy measure	21
22	Kruskal–Wallis	filter to remove cyclical component of time series	22
23	Jonckheere–Terpstra	algorithm for non-linear least-squares regression	23
24	Durbin–Watson	test of normal distribution based on moments	24
25	Hodrick–Prescott	test of equality of two continuous distributions	25
26	Guttman–Kaiser	non-parametric test for several independent samples	26
27	Breusch–Pagan	widely used estimator of survival function	27
28	Nadaraya–Watson	criterion for number of significant components or factors	28
29	Hosmer–Lemeshow	relative entropy measure	29
30	Levenberg–Marquardt	convergence diagnostic for MCMC algorithms	30
31	Newton–Raphson	non-parametric test for several ordered independent samples	31
32	Gelman–Rubin	ranks-based two-sample test of stochastic order	32

The methods are of different scope – tests, statistics, estimates, criteria, procedures; some very specific, some very general. The descriptions are brief and simplistic, but I have tried to give hints for distinguishing between methods of similar scope. Another hint: no method is listed next to its description.

Naturally, this is only one possible selection among many for the given "draw size", and there could be many other draw sizes. I have created the list "off the top of my head", but (real or perceived) "imbalance" may be due to availability of "statistical doubles" in addition to my own ignorance and bias. The only limitation I had to observe was that a single person cannot "play" in more than one double even in a hypothetical tournament (so Box-Jenkins and Cox-Snell cannot be listed together with Box-Cox, for example).

The quiz is limited to gentlemen's doubles (so, e.g., Navratilova–Shriver are not mentioned among the tennis legends). Obviously, ladies' doubles and mixed doubles are fewer in statistical-method authorship, albeit not absent (think of DerSimonian–Laird and Kendall–Babbington-Smith, respectively).

This quiz could perhaps inspire someone to write a historical essay on naming of statistical methods. Or, perhaps more importantly, it might inspire the readers to put statistics aside for a while and play a game of tennis!

Solution: 1-21, 2-29, 3-13, 4-11, 5-20, 6-24, 7-25, 8-32, 9-8, 10-7, 11-5, 12-27, 13-9, 14-17, 15-2, 16-6, 17-3, 18-10, 19-18, 20-16, 21-1, 22-26, 23-31, 24-12, 25-22, 26-28, 27-19, 28-4, 29-15, 30-23, 31-14, 32-30