The colour blind problem

E. Khmaladze

Affiliation: *Victoria University of Wellington, New Zealand*

E-mail: estate.khmaladze@vuw.ac.nz

Consider an experiment when one observes pairs of balls with random diameters $(X_i, Y_i)_{1\le i\le n}$. In each pair, one ball, with diameter X_i, is green, while the other ball, with diameter Y_i, is blue. The pairs $(X_i, Y_i)_{1\le i\le n}$ are i.i.d., and within each pair X_i and Y_i are independent.

We want to test that diameters of green and blue balls have the same distribution. Denoting these distributions by P_1 and P_2, we want, therefore, to test non-parametric hypothesis

$$H_0: \ P_1 = P_2 \text{ (and equal to some unspecified } Q).$$

As it is formulated so far, the problem is a classical one and the class of test statistics is well understood. Namely, if P_{1n} and P_{2n} are empirical distribution functions based on $\{X_i\}_{1\le i\le n}$ and $\{Y_i\}_{1\le i\le n}$, respectively, any statistic based on the empirical process

$$\sqrt{n}(P_{1n} - P_{2n}), \text{ or } \sqrt{n}[P_{1n} - \frac{1}{2}(P_{1n} + P_{2n})],$$

invariant under time transformation $Q(x) = t$, have distribution which is the same for all Q. If we are interested in change only between expected values of diameters of green and blue balls, then just Student's statistic $\sqrt{n}(\bar{X}_n - \bar{Y}_n)$ (with proper normalisation) will provide a good test.

However, what can one do in the case when an observer cannot distinguish the colours of the balls, that is, when he is "colour blind"?

In this case, it is not possible to construct P_{1n} and P_{2n}, or even averages $\bar{X}_n$ and $\bar{Y}_n$. Yet, we show that systematic approach to testing H_0 is possible, again providing an empirical process and test statistics, based on this process, with distribution independent of (unknown) common Q and explains what is the price one pays for "colour blindness" in terms of the power of our tests.

Joint work with Laura Dumitrescu.