Cosine Quantogram

Kendall's 1974 derivation of the cosine quantogram as a method for determining the existence of as
as deriving an estimate for a quantum unit is based on the following:

> Suppose we wish to demonstrate that an unknown quantum unit q has given rise to our data, where Y_i represents the measurement of object i, i=1,…,N. Specifically we hypothesize that the ith observation, Y_i, is an unknown number of integer multiples M_i of our unknown quantum unit plus a small amount of unknown error. Consider for a moment a simple model $Y_i = M_i q + \varepsilon_i$, i = 1, …, N. For a fixed q, if q is a quantum, the associated errors should be small. For q over some range (q_L, q_U), calculate
>
> $$\phi(q) = \sqrt{\frac{2}{N}} \sum_{i=1}^{N} \cos(2\pi x_i / q)$$
>
> Graph $\phi(q)$ vs. q; call this the cosine quantogram.

We wish to find q* such that $\phi(q)$ is maximized. This will be our maximum likelihood estimate. Note that
cosine quantogram is proportional to the log-likelihood for a von Mises distribution with density function

$$f(x \mid \delta, q) = (2\pi I_0(\delta))^{-1} e^{\delta \cos(2\pi x / q)}$$

ere $\delta \geq 0$, x>0, q>0 and $I_0(\delta)$ is a modified Bessel function of order 0. The maximum is determined numerically
some q in the interval (q_L, q_U). Prior to undertaking this numerical procedure, however, Kendall
ommends "unrounding the data". The process of unrounding involves adding a random value within [−0.05,
5) for measurements rounded to the nearest decimeter, a random value within [−0.005, 0.005) for
asurements rounded to the nearest centimeter, and so on.

Approach: A Generalized Framework for Quantal Inference

We have developed a generalized framework for inference for quanta as follows:

> Let Y_{ijkl} represent the measurement of block l from building i of type j with measurement taken along dimension k. In general we can express the model for quantum identification as follows:
>
> $$Y_{ijkl} = M_{ijkl}\, q_{ijk} + \varepsilon_{ijkl},$$
>
> where M_{ijkl} is an unknown positive integer, q_{ijk} is the unknown quantum that we wish to estimate, and ε_{ijkl} is an error term.

We can use this framework now to construct hypotheses regarding the q's, then build test statistics
ed on the ratio of the likelihood with respect to one hypothesis vs. a suitable alternative. Likelihood theory
es that the maxima of the relevant cosine quantograms can be used to construct this test. In particular we
ht wish to test the hypothesis that the quantum units used for all buildings in our sample are the same. We
write this hypothesis symbolically as a null hypothesis $H_0: q_{i..} = q_{...}$, all i, where we use the "dot" notation to
cate that the factor that is represented with '.' is averaged over in the analysis. For inferential purposes we
write our alternative hypothesis as $H_A: q_{i..} \neq q_{...}$, for at least one value of i. We note that our inferential
hod will allow for testing the identity of the quantum units used in 3 or more buildings as well.

Our method then proceeds in stepwise fashion as follows:

Step 1: Unround the analytic dataset.

Step 2: Derive cosine quantogram and numerically estimate the quantum unit. Call this q*.

Step 3: Randomly resample from the empirical distribution function of measurements. Repeat Step 2.

Step 4: Repeat Step 3 a large number of times. Obtain confidence limit estimates from the empirical
ribution function of the estimates obtained.

Step 5: For the unrounded data from Step 1, calculate the likelihood of the data under H_0 and under H_A. e the ratio of these likelihoods (or alternatively, the difference between the log-likelihoods). Call this the test istic.

Step 6: For the random resamples generated in step 4, calculate the likelihoods and the difference ween the log-likelihoods.

Step 7: Obtain the empirical distribution function of these log-likelihoods.

Step 8: From this empirical distribution function, derive a critical region for a test of size α.

Step 9: Determine if the test statistic falls inside or outside of the critical region, and reject or not reject n favor of H_A accordingly.

If we assume that the data follow a von Mises distribution, we use as our test statistic the logarithm of likelihood ratio (LLR) written as follows:

$$LLR = \sum_i \sum_j \sum_k \sum_l \delta \cos(2\pi y_{ijkl} / \hat{q}_{ijk}^{*0}) - \sum_i \sum_j \sum_k \sum_l \delta \cos(2\pi y_{ijkl} / \hat{q}_{ijk}^{*A}) + c(\tilde{N})$$

ere $c(\tilde{N})$ is a constant function dependent only on $\tilde{N}$, the vector of the sample sizes of units for all possible ıbinations of the indices. Here the superscripts *0 and *A are used to denote under which hypothesis the ntal value(s) and resulting error terms are estimated using the cosine quantogram. Let T_1 designate the term in the LLR and T_2 the second term. Since the constant will not matter in our testing, our test statistic $_{diff}=T_1-T_2$. However note that T_{diff} is also a function of the parameter δ. Still, we can factor out δ and use as test statistic T_{diff}/δ without loss of power. We can resample the dataset to arrive at point and interval mates for the q^*_{ijk} as well as to derive test boundaries for T_{diff}/δ. We note that we can develop T_{diff}/δ's for er hypotheses, use the cosine quantograms to estimate the q's and perform our hypothesis test.