"The essence of science is replication: a scientist should always be concerned about what would happen if he or another scientist were to repeat his experiment." (Guttman)

p_{rep}

In 2006, the *Association for Psychological Science* introduced in the "*author guidelines*" of *Psychological Science* a new norm of publication:

> **Statistics**
>
> Effect sizes should accompany major results. **In addition, authors are encouraged to use prep rather than p values** (see the article by Killeen in the May 2005 issue of Psychological Science, Vol. 16, pp. 345-353).

Killeen's p_{rep} (Killeen, 2005a) has routinely appeared in *Psychological Science*. We also found its use in 15 other journals [Web of Science review of articles citing Killeen (2005a), april 24 2008]:

Behavioral and Brain Functions
Cerebrovascular Diseases
Consciousness and Cognition
Developmental Psychology
European Journal of Cognitive Psychology
Evolution and Human Behavior
Human Communication Research
Journal of Experimental Psychology: Applied
Journal of Experimental Psychology: Learning, Memory, and Cognition
Journal of Memory and Language
Journal of Research in Personality
Language and Cognitive Processes
Perception
Psychological science
Psychonomic Bulletin & Review
The Quarterly Journal of Experimental Psychology

It is essentially used associated either with a Student's ***t* test for comparing means** or an **ANOVA *F* test with one degree of freedom in the numerator**. So we will restrict our attention to this situation.

- p_{rep} ("**probability of replication**") is the predictive probability, given the data of the current experiment, to **find again a same-sign effect in a replication** of this experiment. From a practical viewpoint, it can be derived from the observed p value only; consequently, from a formal viewpoint, it is **equivalent to *p***. Of course it has a different interpretation, since it is a **predictive** expression of the statistical result of the experiment.

p_{rep} can be derived either from Fisher's **fiducial** argument as by a **Bayesian** assuming noninformative priors (Killeen, 2005b).

Killeen, P.R. (2005a). An alternative to null-hypothesis significance tests. *Psychological Science*, *16*, 345-353.
Killeen, P.R. (2005b). Replicability, Confidence, and Priors. *Psychological Science*, *16*, 1009-1012.

We have enjoyed constating that for the first time a **"natural" probability** - that is a probability going from the known (the data in hand) to the unknown (observations to come) - was routinely reported in psychological journals.

However, without speaking of other uses of the fiducial-Bayesian probabilities, this practice may be improved, both technically and conceptually.

A careful examination of the articles published in *Psychological Science* revealed us that many authors incorrectly used the available formulae, apparently confusing one-tailed and two-tailed *p* values. This reveals a serious implementation problem.

In **about half articles** published in the october issue for each of the two years 2006 and 2007, $\boldsymbol{p_{rep}}$ **was found to be systematically undervalued**. In the majority of these articles, the reported values could be obtained with the formulae given by Killen if we (erroneously) computed them with the **two-tailed** *p* value (instead of the one-tailed *p* value).

The authors who report p_{rep} **merely add it** to the test statistic and/or the *p* value. One can be afraid that they (and their readers) continue to focus on the statistical significance of the results. This attitude could be reinforced by the fact, strongly suggested by our experimental findings, that p_{rep}, the predictive probability of a same-sign result, could be confused with the predictive probability of a same-sign **and significant** result.

Only a solution that assumes a **known variance** has been implemented and is currently used. More than one hundred years after Student's famous article (Student, 1906), one can hardly be satisfied with this unnecessary restriction.

Relaxing the assumption of known variance, p_{rep} and p_{srep}, the probability of a significant replication at one-tailed level α, can be computed from the predictive distribution of the test statistic (or equivalently from the predictive distribution of Cohen's *d*). If *t*2 denotes the test statistic in the replication, assuming for instance that *t*1, the observed value in the current experiment is positive, $\boldsymbol{p_{rep}}$ **is the probability that** ***t*2 is positive** and $\boldsymbol{p_{srep}}$ **is the probability that** ***t*2 exceeds** $\boldsymbol{t_\alpha}$, the 100α percent upper point of the Student distribution with the same number of degrees of freedom as for the test statistic in the current experiment.

Lecoutre (1984) called the fiducial-Bayesian predictive distribution of the *t* test statistic a ***K*-prime** distribution. This distribution was studied in details in Lecoutre (1999). An algorithm for computing its cumulative distribution function was given in Poitevineau and Lecoutre (2010).

Computing p_{rep} with Excel

In the known variance case, Killeen (2005a) gave the following formula for Excel users:

p_{rep} = **NORMSDIST(NORMSINV(1-p)/SQRT(2))**, where **p** is the **two tailed** *p* value of the *z* test

This formula can be generalized for an unknown variance:

p_{rep} = **1-TDIST(TINV(2*p,df)/SQRT(2),df),1)**

where **p** is the **two tailed** *p* value of the *t* test (for an ANOVA F test, **halve p**) and **df** is the number of degrees of freedom.

p_{rep} can also be directly computed from the test statistic, either Student's t or ANOVA F with one degree of freedom in the numerator:

p_{rep} = **1-TDIST(ABS(t)/SQRT(2),df,1)**

p_{rep} = **1-TDIST(SQRT(F)/SQRT(2),df,1)**

Getting p_{rep} and p_{srep} from tables

A detailed table gives $\boldsymbol{p_{rep}}$ as a function of the two tailed *p* value.

A detailed table that gives $\boldsymbol{p_{srep}}$ (for α=.05) as a function of the two tailed *p* value.

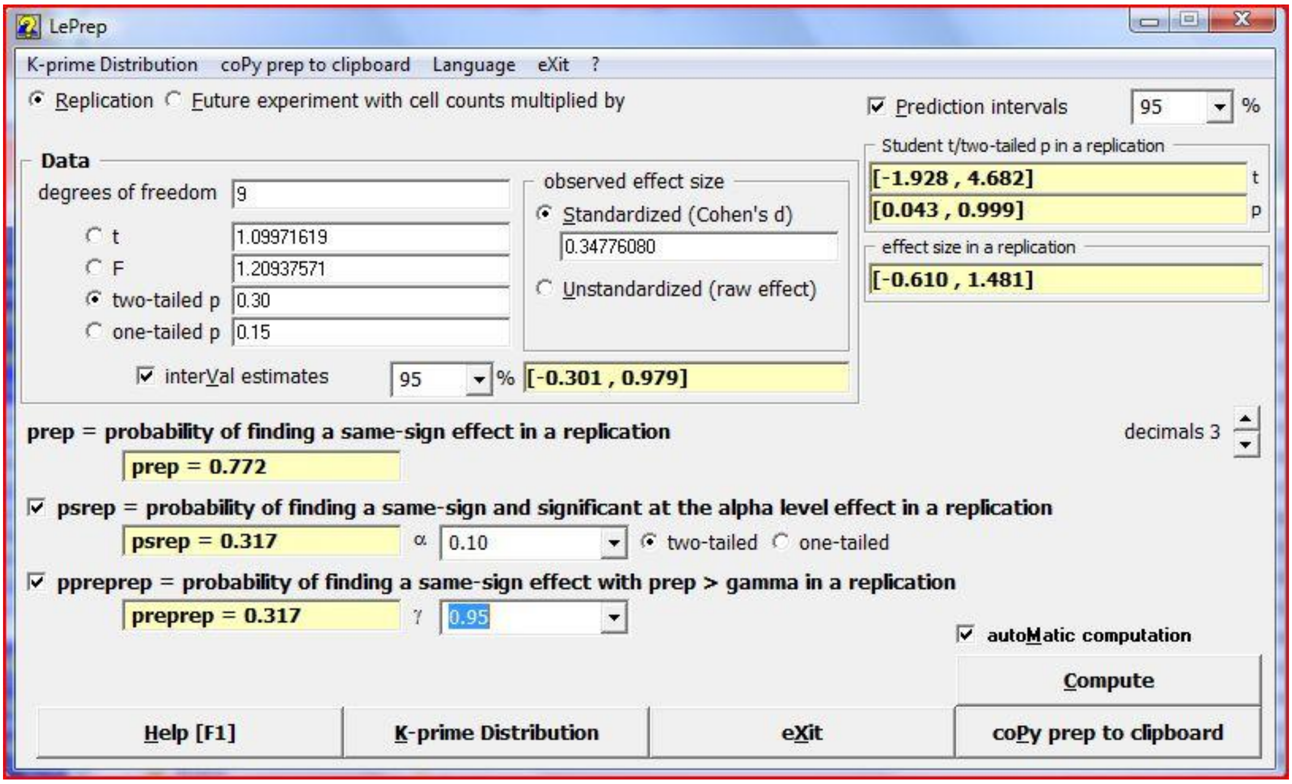
LePrep: a friendly-user Windows program

Compute :

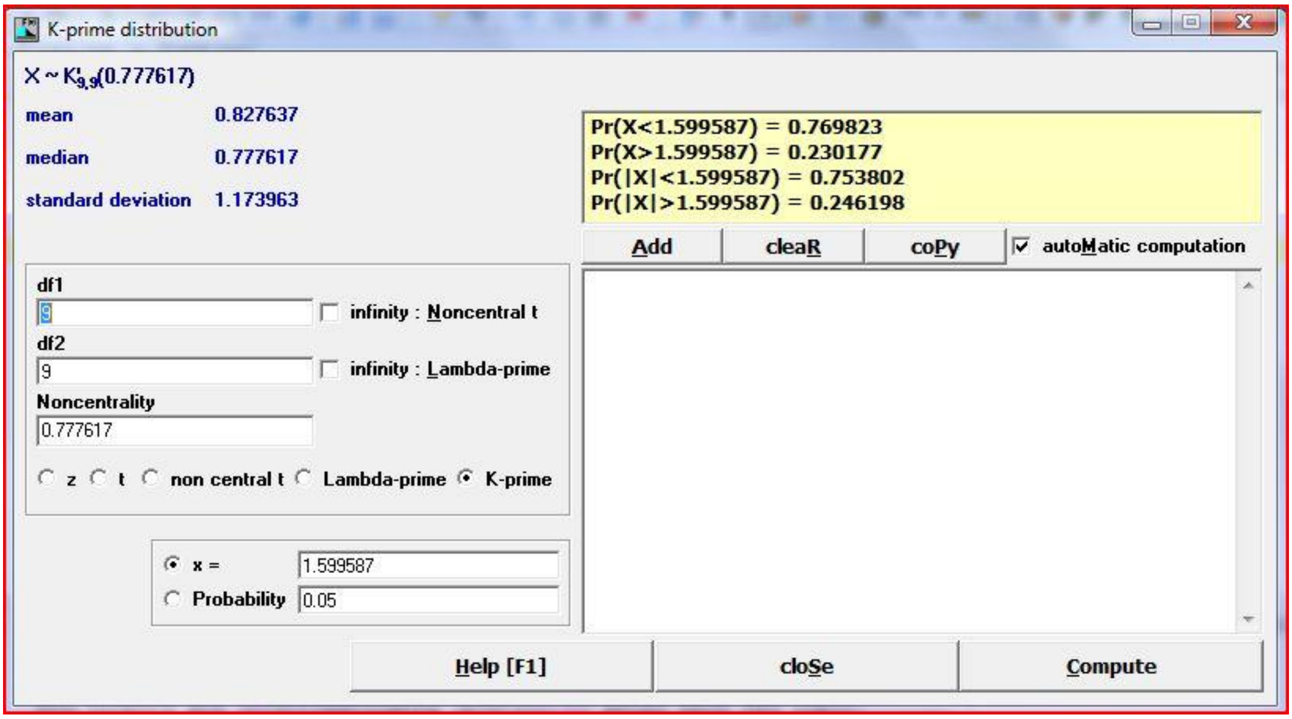
- the predictive probability **prep** ("Killeen's probability of replication") of finding **a same-sign** effect in a replication,
- the predictive probability **psrep** of finding **a same-sign and significant at one-tailed level α** effect in a replication,
- the predictive probability **ppreprep** of finding **a same-sign effect with prep larger than γ** in a replication.

It can also be used as a **Word macro**

Interval estimates for a contrast: standardized or unstandardized



K-prime distribution



References

- **Lecoutre B. (1984) –** *L'Analyse Bayésienne des Comparaisons*. Lille: Presses Universitaires de Lille.
- **Lecoutre B. (1999) –** Two useful distributions for Bayesian predictive procedures under normal models. *Journal of Statistical Planning and Inference*, *77*, 93-105.
- **Poitevineau J., Lecoutre B. (2010) –** Implementing Bayesian predictive procedures: The K-prime and K-square distributions. *Computational Statistics & Data Analysis*, *54*, 723-730.
- **Lecoutre B., Lecoutre M.-P., Poitevineau J. (2010) –** Killeen's probability of replication and predictive probabilities: How to compute, use and interpret them. *Psychological Methods*, *15*, 158-171.
- **Lecoutre B., Killeen P. (2010) –** Replication is not coincidence: Reply to Iverson, Lee, and Wagenmakers. *Psychonomic Bulletin & Review*, *17*, 263-269.