CHANCE MODELS: BUILDING BLOCKS FOR SOUND STATISTICAL REASONING

Herman Callaert

Center for Statistics, Hasselt University, Belgium

herman.callaert@uhasselt.be

A good understanding of chance models is crucial for mastering basic ideas in statistical inference. Mature students should be introduced to the concepts of inference through a study of the underlying chance mechanisms. They should learn to think globally, in models. In an introductory course, these models should have their own clear and unambiguous notation. Fuzziness and flaws, as encountered by our students in textbooks and software, may hamper their learning process seriously. The above claims are based on my experience as an instructor for university students. They should be substantiated by systematic research on the potential advantage of "thinking in models", possibly also for younger pupils.

INTRODUCTION

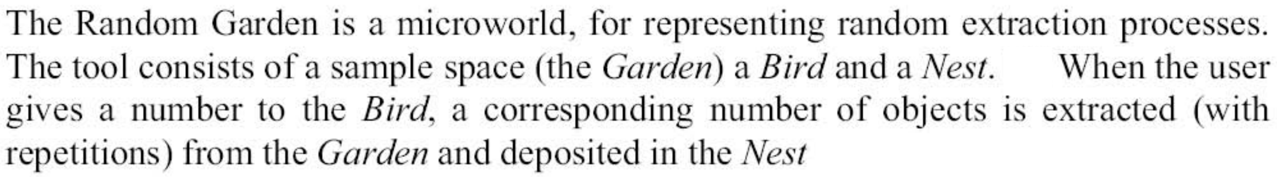
From my experience as a teacher of statistics, thinking in models is a stumbling block for many mature students when they are confronted with the basic concepts of statistical inference. As long as students do not master the connection between underlying chance mechanisms and statistical conclusions, procedures like the construction of confidence intervals remain "black boxes". The main problems with confidence intervals have been discussed in a previous paper (Callaert 2007) where the ability of "thinking backwards" was shown to be essential. After seeing the data, the main question was: "how did those data come to me?" This is a question about an underlying probability model as an ideal mathematical construct for modelling outcomes in a physical world. Those models are the main theme of the current paper.

This paper has two parts. It first shows how mathematical mature students can be introduced to chance models at all places, from populations over samples to statistics. A simple example illustrates how the models are built. It points at the same time to the fact that a clear and unambiguous notation is crucial for acquiring clear and unambiguous insight. Students discover the need for distinguishing a population mean from a sample mean, or an "observable" chance model from an "unobservable" but fixed parameter. Many of the inaccuracies found in research papers, textbooks and software packages have their origin in a lack of insight in underlying chance models. Some examples are given in the second part of this paper.

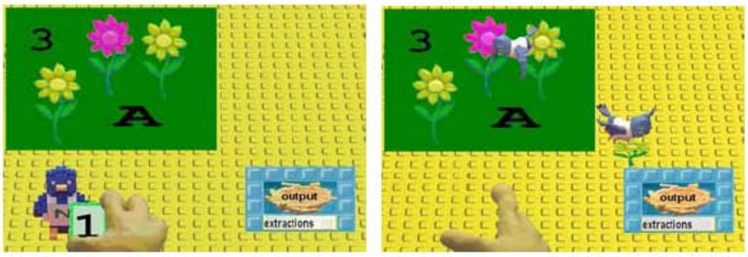
The current text is focused on mathematical mature students (using explicit mathematical notation). The underlying concepts however are very fundamental and it certainly is worthwhile finding out what can be done with younger pupils. Research by Prodromou (2007) and Prodromou and Pratt (2006) is most interesting in this respect. They look at the connection between a *data-centric* and a *modelling* view on distributions, and write that: *“The modelling perspective reflects the mindset of statisticians when applying classical statistical inference”*. How and at what age can the connection with statistical inference be made?

THE POPULATION AS A CHANCE MODEL

From the very start, it is important that pupils not only are interested in “what” comes to them but also in “how” it comes to them. When they are allowed to build their own chance mechanisms, it is clear that (after some time and some experiments) they focus on both aspects. Nice examples can be found in a variety of research papers, such as in the study carried out by Pratt (1998) where children are able to manipulate “the underlying chance mechanism” (workings box). Another example is described in a paper by Cerulli et al. (2007) where they write:



The Random Garden is a microworld, for representing random extraction processes. The tool consists of a sample space (the *Garden*) a *Bird* and a *Nest*. When the user gives a number to the *Bird*, a corresponding number of objects is extracted (with repetitions) from the *Garden* and deposited in the *Nest*



In that study, one team of pupils creates not just a Garden but a *Random Garden*. This means that the pupils not only think about the composition of the garden (the flowers and trees) but they also know that the *Bird* will extract objects “at random and with replacement”. A competing team of pupils has to guess the *Random Garden* after they have inspected a *Nest*. That the objects in the *Nest* came “at random and with replacement” is key information and it is used (rather implicitly) by the competing team when they look at bar graphs and counters. One of the important consequences of the setup of this study is that pupils start discussing (and understanding) the concept of “equivalent chance mechanisms” (called equivalent gardens). If the study would have been set up differently, with the same flowers and trees but with a *Bird* that extracts objects not at random or without replacement, the “*Guess my Garden Game*” would have been completely different. This aspect might be stressed even more in such types of studies since it is important to find out at what age pupils are able to “think in models” and what kind of strategies can be used for enhancing (and evaluating) this type of thought-processes.

The above examples refer to studies with younger pupils (such as 11-12 years old). At a later stage the concrete objects in populations (such as flowers or colored segments) are replaced by numbers. But the basic question about a population stays unaltered: “which numbers will come to me and with what probability?” For mathematical mature students, comfortable with abstract notation, it is helpful to make a distinction between a chance model and its outcomes. In line with the notation for random variables, a chance model can be denoted by a capital letter (such as X) and outcomes by the corresponding small letter x. An example of such a “population chance model” is what I call a *red die*. Physically, it is just a regular die (falling on each side with probability 1/6) but the numbers on the faces have been changed. There are 3 faces with a 1, 2 faces with a 3, and one face with a 6. The way in which outcomes from this population appear is governed by a throw of this red die. Hence, one will never see a number 2 but, for example, one will get a number 3 with probability $2/6$, denoted as $P(X=3)=2/6$. The next table gives complete information about this population X .

x	1	3	6
$P(X=x)$	$\frac{3}{6}$	$\frac{2}{6}$	$\frac{1}{6}$

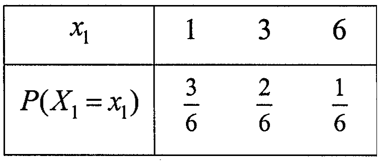
Table 1. The population X described by its chance model

Remark that also in the continuous case it is customary to describe a population by providing at the same time the range of the values and their chance behavior, as reflected by statements like: “we work with a normal $N(124;16)$ population”.

THE SAMPLE AS A CHANCE MODEL

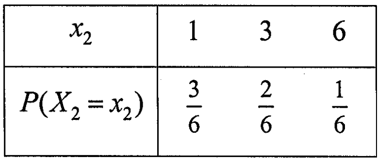
Once students get used to look at populations from the perspective of chance models one would think that the step towards looking at a sample from the same perspective is straightforward. For most of my students, this was not evident. The following (simple) example became a real eye-opener for many of them.

What happens when one takes a sample of size $n=2$ from the population X described in table 1 (the *red die*)? The main point here is that students have to answer the question *before* they actually take the sample. Hence, the question: “What will be the result of the first draw?” is not answered by “How can I know?” (reasoning only about specific outcomes *after* an experiment has been carried out) but by “I can tell you, *beforehand*, every possible value together with its probability”. And then of course it is not difficult to come up with the chance model X_1 for the first draw. The second draw X_2 has the same behavior.



x_1	1	3	6
$P(X_1=x_1)$	$\frac{3}{6}$	$\frac{2}{6}$	$\frac{1}{6}$

Table 2.



x_2	1	3	6
$P(X_2=x_2)$	$\frac{3}{6}$	$\frac{2}{6}$	$\frac{1}{6}$

Table 3.

A model for a sample of size $n = 2$ now follows easily from tables 2 and 3. The model is denoted by (X_1, X_2) and its outcomes by (x_1, x_2). It is instructive for students to construct this model for themselves arriving at table A1 (appendix) or at an urn model with random draws from the urn (figure 1).

The insight that a sample result (x_1, x_2) is nothing but one of the possible outcomes of an underlying chance mechanism (X_1, X_2) is very important. It creates the appropriate context for a proper understanding of the behavior of the sample mean (or of any other statistic constructed from a sample).

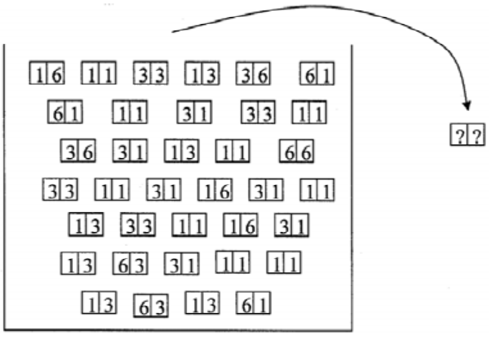


Figure 1.

THE SAMPLE MEAN AS A CHANCE MODEL

Continuing the above example, it takes just a few minutes to find all possible values of the sample mean together with their corresponding probabilities (see table A2 in the appendix). This leads to the following model:

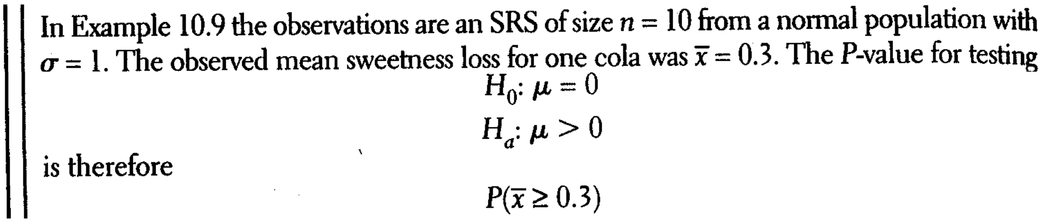
$\bar{x}$	1	2	3	3.5	4.5	6
$P(\bar{X} = \bar{x})$	$\frac{9}{36}$	$\frac{12}{36}$	$\frac{4}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{1}{36}$

Table 4. The sample mean $\bar{X} = \dfrac{X_1 + X_2}{2}$ described by its chance model

Simulation tools might be extremely useful for learning statistical concepts but it is my experience that mature students (and secondary school mathematics teachers) also need an explicit confrontation with the more abstract tool of “thinking in models”. For many of them, the behavior of a sample mean is better understood in the context of chance models like table 4 than through the experience that a simulated bar chart or histogram is an approximation of a so-called sampling distribution. Properties like: “the mean of the sample mean is the population mean” can be discovered through simulations, but a clear view on underlying models surely can enrich insight in this discovery. In either case, an unambiguous notation is needed as a support to students for distinguishing populations from samples, and chance models from their outcomes. The next sections illustrate some problems.

EXAMPLES FROM TEXTBOOKS

During the past couple of decades reform in statistics education at the school level has been extensive in the United States. It has resulted in the production of new textbooks by authors such as: Yates, Moore and Starnes (2003) [YMS], Watkins, Scheaffer and Cobb (2004), Agresti and Franklin (2007), and many others. All these books use capital letters (such as X) for random variables and small letters (such as $x_1, x_2 ...$) for their outcomes. This is nice since this notation makes a clear distinction between an underlying chance process and a particular outcome. But once students start sampling, their attention is drawn to particular outcomes and the notation for underlying models, such as (capital) $\bar{X}$ for the sample mean, is gone. Paul Velleman, author of ActivStats, says: "*Convention in the introductory course is to emphasize the observed values, which are usually not thought of as random. Every text I know uses a lower case $\bar{x}$ to represent the sample mean. The r.v. version is a hypothetical construct of which the sample mean at hand is one realization. A bit sloppy at times, but, I think, less confusing for students*" [(1999) personal communication]. The experience I have with my students tells me the opposite. On p.525 of [YMS] one reads: "The sampling distribution of $\bar{x}$ describes how the statistic $\bar{x}$ varies in all possible samples from the population. The mean of the sampling distribution is μ, so that $\bar{x}$ is an unbiased estimator of μ". The fact that $\bar{x}$ stands for an outcome while at the same time it is said that $\bar{x}$ is unbiased is confusing. The problem persists in the chapter on hypothesis testing where one reads on p.568 that $\bar{x} = 0.3$ and that $P(\bar{x} \geq 0.3)$ is needed for computing the p-value. But probability statements are statements about chance processes. Hence, the p-value is the probability that (under the null hypothesis) the chance process $\bar{X}$ generates values which are at least as large as the observed outcome $\bar{x}$. Notation is crucial here and the above phrase should be written as $P(\bar{X} \geq \bar{x})$. If $\bar{x} = 0.3$ in the sample of one student while $\bar{x} = 0.4$ in the sample of another student, they now can start with the same notation $P(\bar{X} \geq \bar{x})$. Afterwards, they only have to plug in their $\bar{x}$-value for arriving at $P(\bar{X} \geq 0.3)$ [or at $P(\bar{X} \geq 0.4)$] as meaningful expressions.



In Example 10.9 the observations are an SRS of size $n = 10$ from a normal population with $\sigma = 1$. The observed mean sweetness loss for one cola was $\bar{x} = 0.3$. The P-value for testing

$$H_0: \mu = 0$$

$$H_a: \mu > 0$$

is therefore

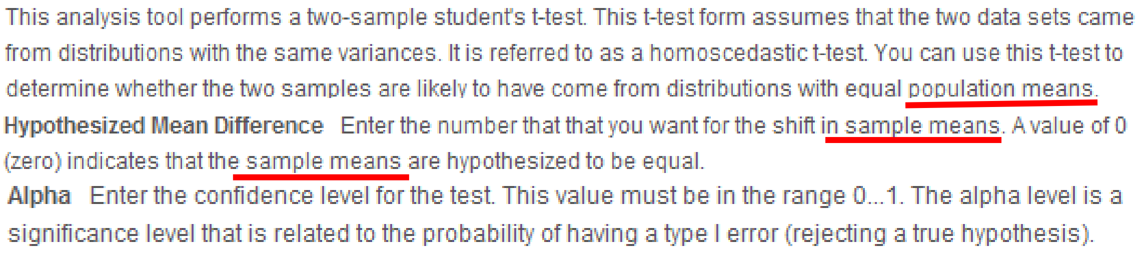
$$P(\bar{x} \geq 0.3)$$

EXAMPLES FROM SOFTWARE

Software can provide powerful educational tools and can create unique opportunities for gaining insight in statistical concepts. This is not only true for our students but also for adults who (sporadically) need to carry out a statistical analysis. At those instances, people often use their favorite package as a fast resource, both for ideas and for computations. From a "statistical literacy" point of view, one would hope that statistical information encountered in widespread packages is clear and accurate.

Excel

When your student says that, in a one-sided two-sample t-test, the null hypothesis assumes that the two means are equal and the alternative hypothesis says that one mean is larger than the other, you might be willing to consider the answer as correct. But when he writes $H_0 : \bar{x} = \bar{y}$ *versus* $H_1 : \bar{x} > \bar{y}$ you can't believe your eyes. In his notation, he tries to find out whether the mean in his first sample is larger than the mean in his second sample $\bar{x} > \bar{y}$ instead of investigating whether the mean of the first population is larger than the mean of the second population $\mu_1 > \mu_2$. This type of confusion has been present in Excel for decades. Several versions in the nineties had in their "Data Analysis Toolpack" a help file called "Learn about the t-test: Two Sample Assuming Equal Variances Analyses". What you could learn was as follows. *"This analysis tool performs a two-sample Student's t-test. This t-test form assumes that the means of both data sets are equal; it is referred to as a homoscedastic t-test. You can use t-tests to determine whether two sample means are equal"*. Apparently, when you have two datasets you can use the Data Analysis Toolpack in Excel for finding out whether $\bar{x}$ equals $\bar{y}$. And you can do so at some alpha level, as follows. *"Enter the confidence level for the test. This value must be in the range 0…1. The alpha level is a significance level related to the probability of having a type I error (rejecting a true hypothesis)"*. There is no clear distinction between a null and an alternative hypothesis (which is *the true hypothesis* to be rejected?) nor is there any reference to underlying populations. This type of fuzziness is disturbing. Attention to these problems has been drawn at several occasions, even in a publication (Callaert 1999). Change however is slow and confused. In Excel 2003 as well as in Excel 2007 it depends on the order in which you call for help. Press F1 (Help), type the phrase Data Analysis and click Search. Then click on Data Analysis and in the new window click on t-Test. The following text appears.

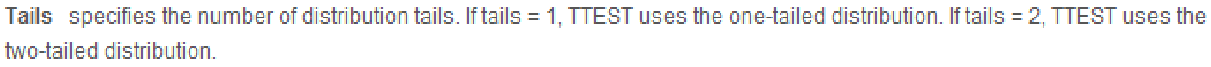


This analysis tool performs a two-sample student's t-test. This t-test form assumes that the two data sets came from distributions with the same variances. It is referred to as a homoscedastic t-test. You can use this t-test to determine whether the two samples are likely to have come from distributions with equal population means.

Hypothesized Mean Difference Enter the number that that you want for the shift in sample means. A value of 0 (zero) indicates that the sample means are hypothesized to be equal.

Alpha Enter the confidence level for the test. This value must be in the range 0...1. The alpha level is a significance level that is related to the probability of having a type I error (rejecting a true hypothesis).

But if you click on Formulas –>More Functions–>Statistical–>TTEST–>"Help on this function", then you can read about equality of *population means* together with a choice of using either a one-tailed or a two-tailed t-distribution.



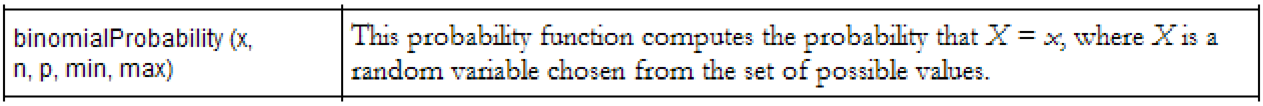
Tails specifies the number of distribution tails. If tails = 1, TTEST uses the one-tailed distribution. If tails = 2, TTEST uses the two-tailed distribution.

Fathom

Never before I've worked with Fathom, so I only can give some first impressions by a novice (having downloaded a Fathom Evaluation Version 2.1). The fact that I was lost right from the start might be blamed on my inexperience. I think however that the rather abstract structure of Fathom working with "collections", "attributes", “measures” and “statistical objects” is not obvious for beginning students. In contrast with this, Maxara and Biehler (2007) report on a study where Fathom was used systematically by their university students, apparently with success. I assume that those students’ first contact with Fathom was different from mine, since I clicked Help–>Sample Documents–>Statistics and started reading. I was quite amazed.

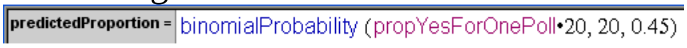
To start with, a clear notation could be helpful. The Fathom Documents use “mu”, “Mean”, “popMean”, “m”, “Avg”,.. and “sigma”, “Std. dev.”, “popSD”, “s”, “sd”,… Why not stick to μ and σ for populations and to $\bar{x}$ and s for sample results?

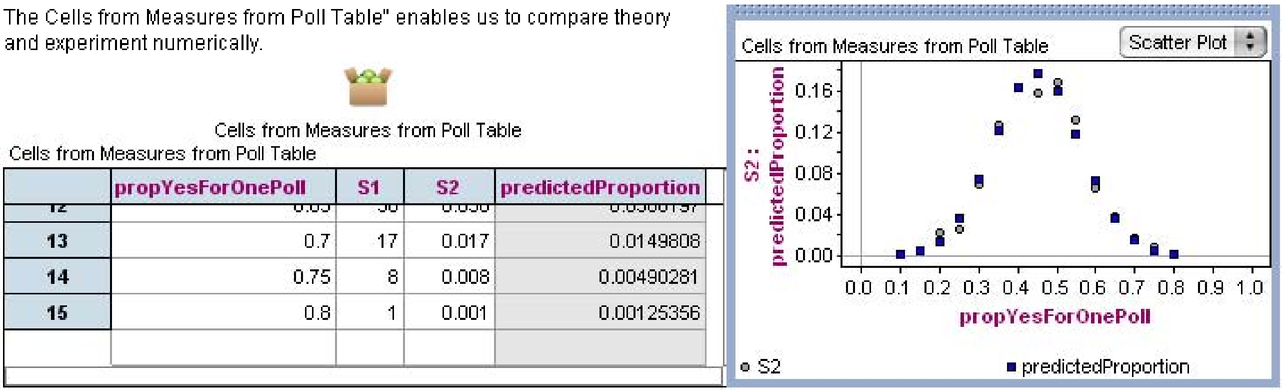
Furthermore, the notational distinction between a binomial model X (capital letter) and its x-values (small letter) should be applauded were it not that X is said to be a random variable chosen from the set of possible values.



binomialProbability (x, n, p, min, max)	This probability function computes the probability that $X = x$, where X is a random variable chosen from the set of possible values.

The binomial model comes up several times but its discrete nature is seldom stressed, even in small samples. The “Polling Simulation” document wants to compare theory and experiment and uses predictedProportion = binomialProbability (propYesForOnePoll•20, 20, 0.45) resulting in a theoretical model where a lot of possible outcomes and their associated probabilities are missing. It is not because one has not seen 17 successes in a particular simulation (and hence not a proportion of 17/20=0.85) that the predicted probability of a proportion of 0.85 doesn’t exist.





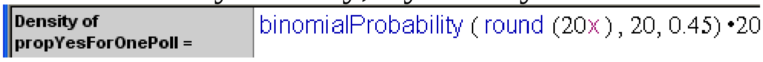
The Cells from Measures from Poll Table" enables us to compare theory and experiment numerically.

Cells from Measures from Poll Table

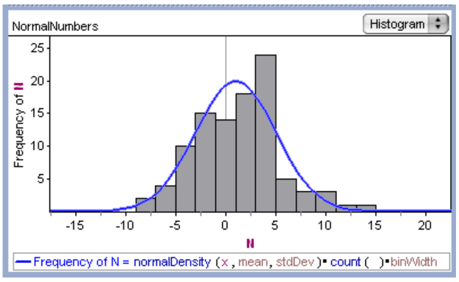
Cells from Measures from Poll Table

	propYesForOnePoll	S1	S2	predictedProportion
12	[illegible]	[illegible]	[illegible]	[illegible]
13	0.7	17	0.017	0.0149808
14	0.75	8	0.008	0.00490281
15	0.8	1	0.001	0.00125356

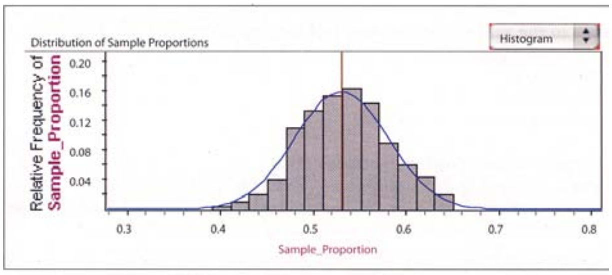
A further problem with this document lies in its histogram representation comparing the simulation results with the (also truncated of course) theoretical model. Repeating a poll of size 20 1000 times does not produce 1000 different outcomes. There still are only 21 different possible proportions. A bar graph comparing theoretical probabilities with experimental relative frequencies would make sense here since the chance model is discrete. By the way, try to let your students discover for themselves the formula Density of propYesForOnePoll = binomialProbability (round (20x), 20, 0.45) •20 for drawing such a histogram. Of course, the problem is much deeper and relates to the obsession of making curves fit histograms who themselves have to represent experiments with discrete outcomes. The “Normal” document for example shows a histogram of 100



random numbers from a normal population together with [quote]: “*a plotted curve of a normal distribution with the same mean, standard deviation, and area as the histogram*”. Yes, with the same area! Fortunately the example uses a histogram on a density scale. But there is no problem if one would use a histogram with frequencies. In the same document under number 3 of the “To do” list attention is drawn to the fact that the density then has to be multiplied by both the count and the bin width. If you do this, you find the figure on the right.

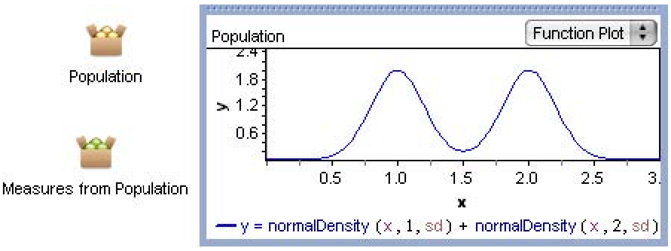


But normalDensity (x, mean, stdDev) • count ( ) •binWidth is a model for what? It is a curve fitting the “frequency histogram” but it certainly isn’t a model for an underlying chance mechanism. These problems are not uncommon. In Schaeffer and Tabor (2008) one finds a similar figure. This time, a histogram has been drawn on a Relative Frequency scale and the density has only been multiplied by the bin width. The authors write: “*The figure shows a simulated sampling distribution of sample proportions. This sampling distribution has a mean of 0.53 and a standard deviation of 0.05 and is nicely represented by the normal distribution (overlaid smooth curve) with that same mean and standard deviation*”. But the top of a normal density $N(0.53;0.05)$ is equal to 8, not to 0.16. So, what’s the name of a bell-shaped curve that (i) is nowhere negative and (ii) has an area under the curve equal to 0.02? Indeed, that’s the blue curve in that paper.



Simulated sampling distribution for a sample proportion

Fathom’s “Central limit Theorem” document has analogous problems. Wouldn’t it be nice to compare the histograms of the simulated sample means $\bar{x}$ with the target model of $\bar{X}$? That model is *normal* with *mean* $\mu = 1.5$ and with *standard deviation* $\sigma/\sqrt{n} = \sqrt{0.5}/\sqrt{n}$. The document instead uses the mean and standard deviation of the randomly generated set of 200 $\bar{x}$-values. Moreover, the collection called “Population” is not the population but contains the sample values, while the population itself is represented by a bimodal curve integrating out to 2 (yes, two).



CONCLUSION

Thinking in chance models might be too abstract for the young learner but at some level in the developmental process the more mature student might need more than “approximations by simulation” in order to fully understand the underlying reasoning of statistical inference. At this point one needs a careful identification of all the involved entities, together with a clear notation, both in textbooks and software. It might be interesting for further research to investigate the impact of an unambiguous notation on the effectiveness of student's learning and understanding of statistics.

REFERENCES

Agresti, A. and Franklin, C.A. (2007). *Statistics: the art and science of learning from data*. Upper Saddle River, NJ: Pearson Education, Inc.

Callaert, H. (1999). Spreadsheets and Statistics: The Formulas and the Words. *Chance* **12**,64.

Callaert, H. (2007). Understanding Confidence Intervals. *Proceedings of the Fifth Conference of the European Society for Research in Mathematics Education.* CD_ROM edited by D. Pitta-Pantazi and G. Philippou.

Cerulli, M., Chioccariello, A. and Lemut, E. (2007). A Microworld to Implant a Germ of Probability. *Proceedings of the Fifth Conference of the European Society for Research in Mathematics Education.* CD_ROM edited by D. Pitta-Pantazi and G. Philippou.

Maxara, C. and Biehler, R. (2007). Constructing Stochastic Simulations with a Computer Tool – Students' Competencies and Difficulties. *Proceedings of the Fifth Conference of the European Society for Research in Mathematics Education.* CD_ROM edited by D. Pitta-Pantazi and G. Philippou.

Pratt, D. (1998). Expressions of Control in Stochastic Processes. *Proceedings of the Fifth International Conference on Teaching of Statistics. Vol 2*. Voorburg (NL). The International Statistical Institute.

Prodromou, T. (2007). Making Connections Between the Two Perspectives on Distribution. *Proceedings of the Fifth Conference of the European Society for Research in Mathematics Education.* CD_ROM edited by D. Pitta-Pantazi and G. Philippou.

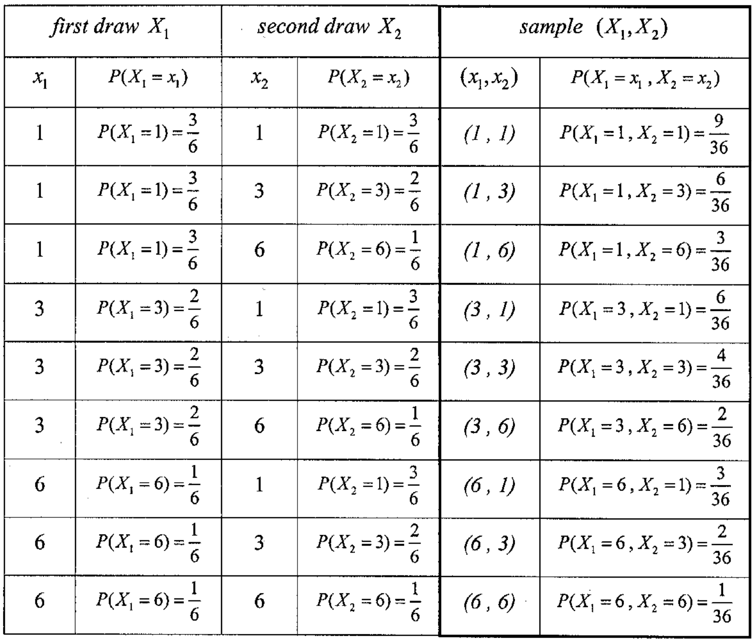
Prodromou, T. and Pratt, D. (2006). The Role of Causality in the Co-ordination of the Two Perspectives on Distribution within a Virtual Simulation. *Statistics Education Research Journal*,5(2),69-88.

Scheaffer, R.L. and Tabor, J. (2008). Statistics in the High School Mathematics Curriculum: Building Sound Reasoning under Uncertain Conditions. *Mathematics Teacher*, **102**, 56–61.

Watkins, A.E., Scheaffer, R.L. and Cobb, G.W. (2004). *Statistics in action. Understanding a world of data.* Emeryville, CA: Key Curriculum Press.

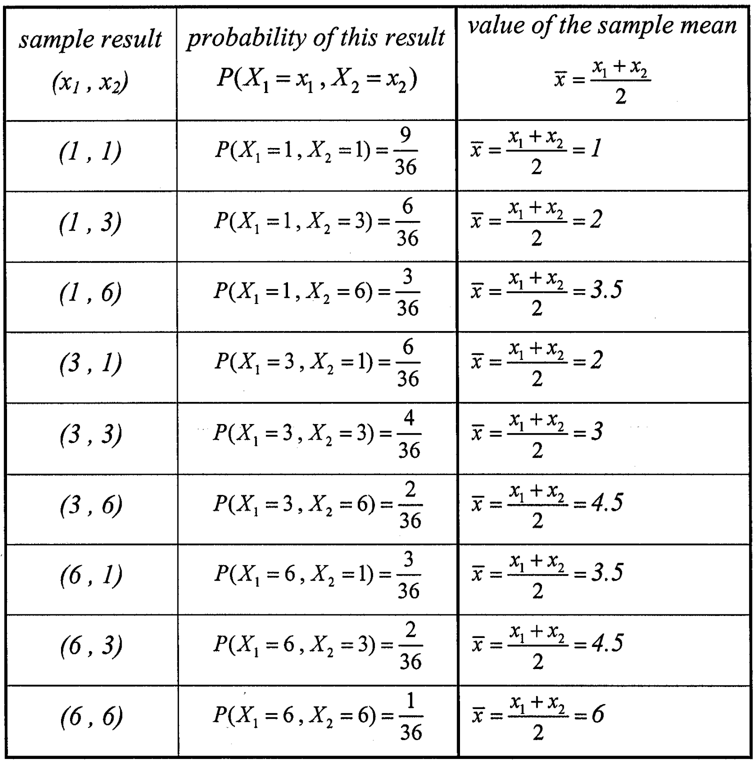
Yates, D.S., Moore, D.S. and Starnes D.S. (2003). *The practice of statistics*. New York: W.H. Freeman and Company.

APPENDIX



first draw X_1		*second draw* X_2		*sample* (X_1, X_2)	
x_1	$P(X_1 = x_1)$	x_2	$P(X_2 = x_2)$	(x_1, x_2)	$P(X_1 = x_1 , X_2 = x_2)$
1	$P(X_1 = 1) = \frac{3}{6}$	1	$P(X_2 = 1) = \frac{3}{6}$	*(1 , 1)*	$P(X_1 = 1, X_2 = 1) = \frac{9}{36}$
1	$P(X_1 = 1) = \frac{3}{6}$	3	$P(X_2 = 3) = \frac{2}{6}$	*(1 , 3)*	$P(X_1 = 1, X_2 = 3) = \frac{6}{36}$
1	$P(X_1 = 1) = \frac{3}{6}$	6	$P(X_2 = 6) = \frac{1}{6}$	*(1 , 6)*	$P(X_1 = 1, X_2 = 6) = \frac{3}{36}$
3	$P(X_1 = 3) = \frac{2}{6}$	1	$P(X_2 = 1) = \frac{3}{6}$	*(3 , 1)*	$P(X_1 = 3, X_2 = 1) = \frac{6}{36}$
3	$P(X_1 = 3) = \frac{2}{6}$	3	$P(X_2 = 3) = \frac{2}{6}$	*(3 , 3)*	$P(X_1 = 3, X_2 = 3) = \frac{4}{36}$
3	$P(X_1 = 3) = \frac{2}{6}$	6	$P(X_2 = 6) = \frac{1}{6}$	*(3 , 6)*	$P(X_1 = 3, X_2 = 6) = \frac{2}{36}$
6	$P(X_1 = 6) = \frac{1}{6}$	1	$P(X_2 = 1) = \frac{3}{6}$	*(6 , 1)*	$P(X_1 = 6, X_2 = 1) = \frac{3}{36}$
6	$P(X_1 = 6) = \frac{1}{6}$	3	$P(X_2 = 3) = \frac{2}{6}$	*(6 , 3)*	$P(X_1 = 6, X_2 = 3) = \frac{2}{36}$
6	$P(X_1 = 6) = \frac{1}{6}$	6	$P(X_2 = 6) = \frac{1}{6}$	*(6 , 6)*	$P(X_1 = 6, X_2 = 6) = \frac{1}{36}$

Table A1. The sample (X_1, X_2) described by its chance model



sample result *(x₁ , x₂)*	*probability of this result* $P(X_1 = x_1 , X_2 = x_2)$	*value of the sample mean* $\bar{x} = \frac{x_1 + x_2}{2}$
(1 , 1)	$P(X_1 = 1, X_2 = 1) = \frac{9}{36}$	$\bar{x} = \frac{x_1 + x_2}{2} = 1$
(1 , 3)	$P(X_1 = 1, X_2 = 3) = \frac{6}{36}$	$\bar{x} = \frac{x_1 + x_2}{2} = 2$
(1 , 6)	$P(X_1 = 1, X_2 = 6) = \frac{3}{36}$	$\bar{x} = \frac{x_1 + x_2}{2} = 3.5$
(3 , 1)	$P(X_1 = 3, X_2 = 1) = \frac{6}{36}$	$\bar{x} = \frac{x_1 + x_2}{2} = 2$
(3 , 3)	$P(X_1 = 3, X_2 = 3) = \frac{4}{36}$	$\bar{x} = \frac{x_1 + x_2}{2} = 3$
(3 , 6)	$P(X_1 = 3, X_2 = 6) = \frac{2}{36}$	$\bar{x} = \frac{x_1 + x_2}{2} = 4.5$
(6 , 1)	$P(X_1 = 6, X_2 = 1) = \frac{3}{36}$	$\bar{x} = \frac{x_1 + x_2}{2} = 3.5$
(6 , 3)	$P(X_1 = 6, X_2 = 3) = \frac{2}{36}$	$\bar{x} = \frac{x_1 + x_2}{2} = 4.5$
(6 , 6)	$P(X_1 = 6, X_2 = 6) = \frac{1}{36}$	$\bar{x} = \frac{x_1 + x_2}{2} = 6$

Table A2. Sample mean values $\bar{x} = \frac{x_1 + x_2}{2}$ for all possible sample outcomes (x_1 , x_2).

The arithmetic mean is computed for all outcomes (x_1 , x_2) from table A1.