Tree Improvement Branch

Germination Tests: How Precise Are They?

How many people have thought of how precise the germination capacity (GC) estimate of a seedlot is? If you determine that a seedlot has a GC of 85% and immediately repeat the test, how close do you think you should be? Many decisions are made based upon a seedlots GC without any regard to the precision or variability in that <u>estimate</u>. To be perfectly clear on precision (some call it efficiency), I mean how consistent or how much variation is present in repeated measurements. In seed testing we generally use multiple replications (usually 4 replications of 100 seeds) and average the results of the 4 replications to obtain a more precise estimate of the GC.

The standard deviation conveys the degree of variation in GC among the replicates. If we wish to provide an estimate of the precision of an overall mean (of the 4 replicates) the standard error is appropriate (Zar 1974). Most basic statistics textbooks can provide more details on these statistics. The sample standard error of a mean is estimated as the standard deviation of the four replicates divided by the square root of the sample size. Another way to think of the standard error is the variation in the estimate (i.e. overall mean GC) after repeating the same sampling process and point estimation over and over again.

This article will focus on the **precision of germination tests**, but more importantly open up the discussion of **what format would be most useful in its presentation**. The intent is to be able to allow for some quantification of the precision of a GC value on an individual seedlot basis. In Table 1, the results of five *Tsuga heterophylla* germination tests (not extremes) are presented to illustrate germination test precision. The four individual replicates are averaged to produce the GC and the standard error is calculated as the standard deviation divided by 2 (square root of 4). Knowing the standard error and the t-value (3.182 in all our examples here with 4 replicates) the 95% confidence intervals are calculated. These indicate the range of values between which we are 95% confident the true population mean falls based on sample data provided.

Seedlot A and B both have a GC of 84.8%, but much greater variation is present in seedlot B with a much wider (less precise) confidence interval. One may think that replicate 2 is abnormal and should be removed to improve the precision of estimate. The International Seed Testing Association (ISTA 1999) does have tolerances for the maximum tolerated range between replicates, but the replicates in this example all fall within that range. We therefore have a less precise estimate of GC in seedlot B. Is there a practical need to quantify this?

In seedlot C the GC is 95.2% and a nursery may single-sow this seed in containers. If the germination is actually 92% will this compromise the crop? Seedlot D is a fairly poor seedlot with a confidence interval that spans over 20 percentage points. How do we integrate this large variability in our sowing decisions? Probability models have generally been used to determine sowing requirements, but none to my knowledge include the GC precision as an input variable.

Table 1. Estimated mean germination capacity (GC), estimated standard error, individual replicate data and 95% confidence interval for five individual seedlots of *Tsuga heterophylla*.

Seedlot	Rep 1	Rep 2	Rep 3	Rep 4	Mean GC	Standard Error	95 %Confidence Interval
A	84	84	88	83	**84.8**	1.11	82.4 → 87.2
B	87	77	88	87	**84.8**	2.59	76.6 → 93.0
C	98	96	94	93	**95.2**	1.11	91.7 → 98.9
D	79	73	64	77	**73.2**	3.33	62.6 → 83.8
E	90	93	88	86	**89.2**	1.49	84.5 → 93.9

There is a large range in the observed precision among the seedlot tests. Are there similar differences at the species level? The <u>average</u> statistics have been generated for each of the species with the most commonly used germination test (Table 2). The average mean and average standard error represent what would be expected after a single test of 4 replicates. This corresponds to the variation among all of the test means for each species.

Table 2. Precision of germination tests by species including sample size, average estimated mean germination capacity (GC), average estimated standard error and average 95% confidence limits.

Species	# Tests (# seedlots)	Average Mean GC	Average Standard Error	Average 95% Confidence Interval
Abies amabilis	**474** (236)	63.1	2.16	56.2 → 70.0
Abies grandis	**137** (65)	66.7	2.05	60.2 → 73.2
Abies lasiocarpa	**383** (188)	56.2	1.98	49.9 → 62.5
Abies procera	**38** (20)	63.2	2.13	56.4 → 70.0
Thuja plicata	**1058** (400)	73.1	1.93	67.0 →79.3
Pseudotsuga menziesii var.menziesii	**583** (372)	90.0	1.39	85.6 → 94.5
Pseudotsuga menziesii var.glauca	**1091** (615)	87.7	1.48	83.0 → 92.5
Tsuga mertensiana	**129** (50)	86.9	1.64	81.7 → 92.1
Tsuga heterophylla	**931** (365)	78.3	1.81	72.6 → 84.1
Larix occidentalis	**556** (201)	77.2	1.80	71.4 → 82.9
Pinus contorta var. contorta	**120** (66)	90.8	1.34	86.5 → 95.0
Pinus contorta var. contorta	**3173** (1798)	92.0	1.20	88.2 → 95.9
Pinus monticola	**251** (158)	82.4	1.79	76.7 → 88.1
Pinus ponderosa	**437** (224)	86.4	1.52	81.6 → 91.3
Picea sitchensis	**322** (197)	90.3	1.35	86.0 → 94.6
Picea glauca / engelmannii complex	**2332** (1225)	82.0	1.76	76.4 → 87.7
Picea lutzii	**124** (50)	84.6	1.67	79.3 → 89.9
Chamaecyparis nootkatensis	**138** (95)	36.8	2.17	29.9 → 43.7

The mean germination capacity and mean standard error are the average of individual germination test results. The number of seedlots is provided as a reference to indicate the number of distinct genetic populations. In discussing species the mean standard error can be used as a good reference point in comparing species as all germination tests contained four replicates. The two *Pinus contorta* varieties displayed the greatest precision and this comes as no surprise to those familiar with the species. What might be surprising is that the *contorta* variety with an average GC of 92% can be expected to have a true seedlot mean of between 88.2% and 95.9% with 95% confidence. At the other end of the spectrum one has *Chamaecyparis nootkatensis* with a GC of 36.8% and a confidence interval of between 29.9% and 43.7%. These values give some sense of species differences, but generally the individual seedlot will be the unit to investigate and quantify germination test precision.

Discussion

An investigation of germination test precision is something that should be addressed. I am too often amazed at how people assume that there is no error in the estimate of a seedlots GC. I am also amazed to hear that the same confidence would be placed in a *Pinus contorta* GC and an *Abies* spp. GC. I am not sure if a quantification of germination test precision is something useful to nurseries, but all sources of variability should be explored as we fine-tune our seedling production systems. There has also been international interest in this subject as ISTA begins to address the question of uncertainty of measurement in all tests.

I have advocated the use of the standard error as it indicates the precision of a germination test. The standard error can also be used to calculate confidence intervals for estimated mean GC values. The standard 95% probability has been used, although one may be content with 90% confidence that the GC will fall within a narrower range? Knowing the standard error one can calculate the confidence interval for any probability. If one wants a higher confidence in the GC estimate the size of the confidence interval will increase. Is there another way to increase our confidence? The simple, not so cost efficient, solution is to simply increase the number of replicates in our sample. For highly valuable crops, with appreciable genetic variation, this may be a realistic solution.

This is my first attempt at looking at the precision of germination testing, but not my last. I would appreciate feedback on the validity and practicality of the methods examined. There are certainly some issues and questions that I will continue to investigate, but this is the current status. Some of these issues are the normality of the GC estimates within a species distribution. Do transformations improve or change the results? Are our newer germination tests showing greater precision? How do unstratified and stratified germination tests of the same seedlot compare in precision? Are seed orchard crops showing more precision or does the recombination that supposedly gives us more genetic variation also give us less precision in our germination tests? There are still many questions to be asked – stay tuned.

References

International Seed Testing Association. 1999. International rules for seed testing rules 1999. Seed Science & Technology 27 supplement.

Zar, J.H. 1974. Biostatistical Analysis. Prentice-Hall, Inc., Englewood Hills, N.J.620 pp.

David Kolotelo, RPF   **Cone and Seed Improvement Officer**
Dave.Kolotelo@gems7.gov.bc.ca   **(604) 541-1683 extension 228**