The Effect of Rounding on the Probability Distribution of Regrading in the U.S. Peanut Industry

Edgar F. Pebe Diaz, B. Wade Brorsen, Kim B. Anderson, Francisca G.-C. Richter, and Phil Kenkel

This article determines the effect of rounding (pointing-off) of grade percentages to the nearest whole number on the probability distribution of regrading in the peanut industry. Results show that rounding causes graders to have to regrade an extra 4% of samples even when they follow all directions and make no mistakes. When rounding was not used, the sample weight had little effect on the probability of regrading. With rounding, the probability of regrading was reduced by beginning with a larger than 500-gram sample. Thus, rounding provides an incentive to take overweight samples in order to avoid regrading. Overweight samples can over-estimate the value of peanuts. A low-cost way to improve peanut grading accuracy would be to round to tenths rather than whole numbers.

Key Words: grading, normal-jump distribution, peanuts, regrading, rounding

Under the U.S. peanut grading system, all loads of peanuts are officially inspected and graded by the U.S. Department of Agriculture's (USDA's) Federal/State Inspection Service (FSIS). Producers or sellers take their loads to the buying points. Here, peanut contract graders grade the loads. The FSIS employs approximately 2,000 graders at about 500 buying points across the producing areas to grade peanuts during harvest from August to November (Dowell, Meyer, and Konstance, 1994).

Accurate pricing of peanuts depends on accurate grading. One possible source of inaccuracy in peanut grading is the practice of some graders who have been observed taking samples slightly greater than the prescribed 500g, presumably to reduce chances of regrading (Anderson, 1998). Due to time constraints and pressure during rush hours of the grading season, graders may use an overweight sample to ensure the allowable tolerance is met if some of the sample weight is lost. For example, if a 500g sample is required, graders may begin with a 501g sample. With

overweight samples, graders tend to overestimate the grade factors measured, and thus assign peanuts a higher price than is merited (Tsai et al., 1993; Whitaker, Dickens, and Giesbretch, 1991; Whitaker et al., 1994).

Regardless of the actual sample weight, graders calculate grade percentages as if the initial sample weight were exactly 500g and round off the percentages to the nearest whole number, as required by the current grading procedure (USDA/FSIS, 1996). Such severe rounding no longer seems necessary.

Although gains can be obtained by improving the grading procedure, there is no published literature on the rounding of peanut grade factors. As shown by the results of this analysis, rounding creates an incentive for graders to use overweight samples in order to avoid regrading. The implications of this study may be useful for government policy, when revising the USDA's *Farmers' Stock Peanuts Inspection Instructions Handbook* (USDA/FSIS, 1996), and for formal training programs for peanut graders. To implement a change in policy, the benefits of a change must be demonstrated. Toward that end, this analysis was undertaken at the request of people in the peanut industry who disagree with the current rounding procedure.

The purpose of this study is to determine the effects of rounding on the probability of regrading. A designed experiment is developed to generate data. Nonparametric and parametric methods are then applied. The nonparametric approach uses the empirical probability distribution function (pdf). Under the parametric approach, a probability distribution function called the normal-jump distribution is used to model weight discrepancy. The probability of regrading with and without rounding is then obtained with Monte Carlo methods.

The Regrading of Peanuts

With the current peanut grading procedure, the measurement of grade factors begins with a sample weight of 500g for truckloads of 10 tons or less. For single loads of over 10 tons, a sample of 1,000g is used (Oklahoma FSIS, 1995). The grade factors measured with the initial sample weight are sound mature kernels (SMK), sound splits (SS), other kernels (OK), total damage (TD), and hulls (H).

Measuring grade factors involves randomness. When measured in tenths of grams, regardless of the grader's ability, there will be a discrepancy between the weight before and after grading. This discrepancy is defined by the difference between the initial sample weight and the total weight of graded material.

The weight discrepancy found after grading comes from the sample weight lost as dust and kernels during the analysis, and infrequent human errors. The weight lost is mostly dust or dirt created when the sample is shelled. Also, small pods or kernels can fall through the sheller grate or get stuck in the grading screen. Under time pressure, graders may neglect to clean the pan containing kernels from a previous analysis, or might accidentally drop some kernels into the pan when grading. These kernels will show up in the next sample graded. Hence, graders sometimes have total kernels and hulls greater than the initial sample weight. Human or mental errors include

errors in recording weights, calculating the percentages, and transcribing the results (Powell, Sheppard, and Dowell, 1994). Usually graders do not lose pods and kernels or make errors, but when they do, the errors may be large.

To check the accuracy of the grade factors measured, graders first round the percentages in each of the five categories to the nearest whole number. If the sum of these percentages is less than 99% or greater than 101%, peanuts must be regraded. Paradoxically, rounding alone can be responsible for the necessity to regrade—even when the grader has made no errors.

Theory and Models

Let S be the sample weight and G_i the real weight of the ith grade factor, $i = 1, ..., 5$, such that $S = \Sigma_{i=1}^{5} G_i$. The measurements used to determine if the sample must be regraded include the following five steps:

- STEP 1: *Weighing each grade factor.* Each of the grade factors is weighed, so $\tilde{G}_i = G_i + w_i$ ($i = 1, ..., 5$). Here, w_i is the weight difference due to dust and kernels, as well as infrequent human errors. $\tilde{G}_i$ is the measured weight for the ith grade factor. Units are grams.

- STEP 2: *Taking percentages.* Regardless of the sample weight, graders take percentages of grade factors with respect to an initial sample of 500g as follows: $P\tilde{G}_i = \tilde{G}_i \times (100/500) = \tilde{G}_i/5$.

- STEP 3: *Rounding percentages to the closest whole number.* Rounding implies adding a value, say u_i, between -0.5 and 0.5 to each grade percentage, such that the sum is the integer closest to the percentage obtained at Step 2. We obtain: $\tilde{G}_i/5 + u_i$.

- STEP 4: *Adding all rounded grade percentages.* The rounded grade percentages are summed by:

(1)
$$T = \sum_{i=1}^{5} (\tilde{G}_i/5 + u_i) = \frac{1}{5} \sum_{i=1}^{5} G_i + \frac{1}{5} \sum_{i=1}^{5} w_i + \sum_{i=1}^{5} u_i.$$

Since the sum of true grade factors is the true sample weight S, expression (1) can be simplified as:

(2)
$$T = S/5 + w/5 + u,$$

where w is the sum of the weight discrepancies, and u is the sum of all rounding errors.

- STEP 5: *Deciding whether to regrade or not.* If T falls outside the interval (99, 101), regrade the sample. Thus, the probability of regrading is the probability of T falling outside of this interval.

To model this procedure following expression (2), we need to fit a probability distribution for T. The total weight discrepancy, w, is modeled as a normal-jump random variable. Widrow, Kollár, and Liu (1996) show it is possible to model the u_i's as independent uniform random variables, each with parameters -0.5 and 0.5. Furthermore, Widrow, Kollár, and Liu demonstrate that under certain conditions, the u_i's are independent of the nonrounded variables. Thus, total weight discrepancy w is independent of the sum of the u_i's, and the total rounding error, u.

However, to confirm rounding does in fact increase the probability of regrading, it is sufficient to recognize that the variable u, due to rounding, adds more variability to T, or increases its variance. So the probability distribution of T will be heavier in the tails than the distribution of $T - u$, and thus the probability of falling in the tails will be higher for T than for $T - u$: $P[99 \le T \le 101] < P[99 \le T - u \le 101]$. Therefore, rounding is shown to increase the probability of regrading.

Nevertheless, it is common knowledge that rounding introduces error in any type of measurement. Our objective is to measure by how much the rounding procedure increases the probability of regrading. For this purpose, we develop two models in the following section.

Models

The problem is approached with two models. The first, based on expression (2), uses an approximation to rounding, but requires only a few estimation procedures. We call it the theoretical model. The second model relies on empirical models for each grade factor, but does not approximate rounding. We designate it the empirical model. Both models lead to similar results.

The Theoretical Model

Recall expression (2): $T = S/5 + w/5 + u$. Let us model T as a random variable consisting of a constant ($S/5$), plus a sum of independent random variables. The weight discrepancy percentage ($w/5$) is a normal-jump variable, and u is modeled as the sum of five independent uniforms ($-0.5, 0.5$). This model produces an estimate, denoted by $\tilde{T}$, which takes continuous values, while the true T is discrete. Therefore, the following equivalency holds:

$$(3) \qquad \hat{P}(T = a) = P\big(\tilde{T} \in [a - 0.5, \ a + 0.5)\big), \qquad a \in Z.$$

Equation (3) states that for any integer a, the probability of T being a can be estimated by the probability of $\tilde{T}$ being in an interval closed in $a - 0.5$ and open in $a + 0.5$. Thus, the probability of regrading is estimated as follows:

$$(4) \quad \hat{P}(\text{regrading}) = 1 - \hat{P}\big(T = [99, \ 101]\big) = 1 - P\big(\tilde{T} \in [99 - 0.5, \ 101 + 0.5)\big)$$
$$= 1 - P\big(98.5 \le \tilde{T} < 101.5\big)$$
$$= 1 - P\big(98.5 - S/5 \le w/5 + u < 101.5 - S/5\big).$$

If parameters for the normal jump are estimated, we can calculate this probability with a Monte Carlo integration. To estimate the probability of regrading without rounding, omit u and proceed as before.

The Empirical Model

Recall the first part of expression (1): $T = \Sigma_{i=1}^{5} (\tilde{G}_i/5 + u_i)$. This is equivalent to:

$$(5) \qquad\qquad T = \sum_{i=1}^{5} \text{round}(\tilde{G}_i/5) = \mathbf{j}' \text{round}(\tilde{\mathbf{G}}/5).$$

Therefore, T can be expressed as a function of the vector of the five grade factors $\tilde{\mathbf{G}}$, where $\mathbf{j}$ is a column vector of ones. The probability of regrading can be written as:

$$(6) \qquad\qquad \hat{P}(\text{regrading}) = 1 - \int I_{[99,101]}\big[\mathbf{j}' \text{round}(\tilde{\mathbf{G}}/5)\big] f(\tilde{\mathbf{G}}) \, d\tilde{\mathbf{G}},$$

where $I_A[a]$ is the indicator function. Equation (6) thus implies integrating the pdf of $\tilde{\mathbf{G}}$ over the range where T is between 99 and 101. This is, in fact, the probability of not having to regrade. One minus this integral gives the probability of regrading. To solve (6), the pdf of $\tilde{\mathbf{G}}$ must be estimated. By the law of total probability, we have:

$$(7) \qquad\qquad f(\tilde{\mathbf{G}}) = \int f(\tilde{\mathbf{G}}/S, W) f(W; \boldsymbol{\theta}) \, dW.$$

So, if we can fit a conditional distribution for each percentage grade factor, *given* the weight discrepancy, and if the distribution of total weight discrepancy is taken as jump-normal, we can estimate the pdf of the vector of grade factors. By definition, each grade factor is a function of the sample weight, and the weight discrepancy (refer to Step 1 of the grading procedure). Therefore, the following set of equations is estimated:

$$(8) \qquad\qquad \tilde{G}_i = g(S, W), \qquad i = 1, ..., 4;$$

$$\tilde{G}_5 = S - W - \sum_{i=1}^{4} \tilde{G}_i.$$

These equations estimate the mean and variance of $\tilde{\mathbf{G}}$ given W. Assuming the vector is normally distributed, we can obtain its estimated density function, replace in (7) to obtain the unconditional density function of $\tilde{\mathbf{G}}$, and finally replace in (6) to estimate the probability of regrading.

Data and Methods

The Data

Because of the lack of data on initial sample weights, and consequently on weight discrepancy, there was a need to develop a designed experiment. Grade factors were

measured by a team of three professional graders and two aides. They were asked to complete the measurements quickly to simulate conditions during the peak of the grading season.[1]

The experiment used eighty-three 2,100g samples of peanuts. Twenty-five samples were runner-type peanuts and 58 samples were Spanish-type peanuts. The samples included a range of quality and cleanliness characteristics. A peanut divider was used to separate the 2,100g sample into four approximately equal subsamples. The four subsamples were then used to create samples of 500, 501.4, 503.7, and 505 grams in order to have four different sizes[2] from the same sample. All grade factors were calculated in grams and percentages with and without rounding.

Estimation of the Grade Factor Equations

For the empirical method, it was necessary to estimate the grade factor equations. These were estimated using random-effects models (fixed-effects models yield very similar results):

$$(9) \qquad \tilde{G}_{ijk} = \beta_0 + \beta_1 S_{jk} + \beta_2 W_{jk} + \upsilon_{ij} + \nu_{ijk}, \qquad i = 1, ..., 4,$$

where $\tilde{G}_{ijk}$ is the weight of the ith factor in grams measured from the jth truckload and kth subsample size, S_{jk} is the initial sample weight in grams, W_{jk} is the weight discrepancy in grams, υ_{ij} is the random component due to truckload or across truckloads, and ν_{ijk} is the error term due to within-truckload variation. A random-effects model was chosen over a fixed-effects model since our data represent only a sample of truckloads from the population. We are more interested in accounting for the variation due to truckloads than in the particular effects of each truckload in our sample. Parameters in equation (9) were estimated with PROC MIXED in SAS (SAS Institute, Inc., 1994).

The Normal-Jump Distribution to Model Weight Discrepancy

Weight discrepancy (w) is assumed to combine two sources of errors. One source is associated with the weight lost as dust and kernels during grading, and the other is attributable to infrequent human errors. Given these specific characteristics, a probability distribution we denote the "normal-jump distribution" was used to model the distribution of weight discrepancy. This distribution is a modification of the mixed diffusion-jump process used in earlier studies to model asset prices (Steigert and Brorsen, 1996) and exchange rate movements (Akgiray and Booth, 1986, 1988;

---

[1] The simulated conditions were such that the graders likely made more errors than usual. Since this is only one team of graders, the estimates of the probability of a human error cannot be extrapolated to the population of all graders. But the results are sufficient to support the use of the normal-jump distribution and to demonstrate that rounding increases the probability of regrading.

[2] An unshelled peanut with only one kernel will weigh about one gram. It is impractical to obtain weights exact to one-tenth of a gram. Therefore, actual weights varied slightly from the target sample sizes.

Oldfield, Rogalski, and Jarrow, 1977; Tucker and Pond, 1988). The two major differences reflected in the normal-jump distribution model are: (*a*) a normal distribution is used rather than a lognormal, and (*b*) this model calls for a single random variable rather than a continuous-time diffusion process.

The normal-jump distribution combines a normally distributed process and a jump process. We assume the process associated with weight lost as dust and kernels follows a normal distribution with mean α and variance σ_N^2. The magnitude of infrequent human errors is also normal, but errors occur in a discrete manner, and so are modeled as a Poisson process.

Thus, the size of the jump or error is normally distributed with mean μ and variance σ_J^2, while the jump intensity or rate of occurrence is $\lambda \geq 0$. The jump process is assumed to be independent of the normal process. The distribution is skewed if μ is not zero and the direction of skewness is the same as the sign of μ. The distribution is leptokurtic since λ is greater than zero.

Total weight discrepancy w is therefore represented with the random variable W, whose normal-jump pdf is given by:

$$(10) \quad f(W; \boldsymbol{\theta}) = \sum_{n=0}^{\infty} \left[\frac{\exp(-\lambda)}{n!} \lambda^n \right] \left[\frac{\exp\left(-[W - (\alpha + n\mu)]^2 / 2\left(\sigma_N^2 + n\sigma_J^2\right)\right)}{\sqrt{2\pi\left(\sigma_N^2 + n\sigma_J^2\right)}} \right],$$

where $\boldsymbol{\theta}$ is the vector of five parameters of the normal-jump distribution. The terms in the summation tend to zero as n increases. For estimation purposes, the sum is calculated up to $n = 7$ (larger values of n caused an underflow). Integration of $f(W; \boldsymbol{\theta})$ to one was verified using the MAPLE 5.3 mathematical program (Char et al., 1993). The function $f(W; \boldsymbol{\theta})$ is nonnegative since each of the terms is nonnegative. Therefore, $f(W; \boldsymbol{\theta})$ is a valid pdf.

The parameters for the distribution of w were estimated maximizing the following log-likelihood function, derived from equation (10), with respect to the five parameters (α, σ_N^2, λ, μ, and σ_J^2):

$$(11) \qquad \text{Ln}(\Omega) = \sum_{t=1}^{n} \ln \left[\sum_{n=0}^{\infty} \frac{1}{\sqrt{2\pi\left(\sigma_N^2 + n\sigma_J^2\right)}} \left[\left(\frac{\exp(-\lambda)}{n!} \right) \lambda^n \right] \right.$$
$$\left. \times \exp\left[-\frac{\left[w_k - (\alpha + n\mu)\right]^2}{2\left(\sigma_N^2 + n\sigma_J^2\right)} \right] \right].$$

The estimation of this nonlinear model requires a numerical optimization algorithm. The parameters are estimated using the NL command and the LOGDEN option in SHAZAM (White, 1993). The numerical Hessian matrix is computed using the NUMCOV option to calculate standard errors.

Likelihood odds ratios were used to compare the fit of the proposed normal-jump distribution versus normal and Student-*t* distributions. Likelihood-ratio tests compare

Table 1. Descriptive Statistics of Four Pricing Methods Using Paired Differences on Peanut Producer Prices

Description	Mean	Std. Dev.	Minimum	Maximum
500g Sample:				
Prices with Rounding	601.43	20.83	547.42	671.12
Prices with No Rounding	601.17	20.29	550.07	665.36
Paired Differences [a]	0.25	3.30	−6.44	7.98
Mean Squared Error	10.94	13.00	0.00	63.70
Number of Observations = 314				
Actual Sample:				
Prices with Rounding	598.81	20.74	547.42	663.34
Prices with No Rounding	598.46	20.24	546.57	660.54
Paired Differences [b]	0.36	3.35	−7.03	8.31
Mean Squared Error	11.34	13.18	0.00	69.04
Number of Observations = 296				

Notes: All prices related to observations requiring regrading have been excluded. Variability of prices is defined as the standard deviation of the paired differences. Producer prices are expressed in dollars. Total number of observations is 332.

[a] The observed value of the *t*-statistic is 1.34, which is less than 1.96 (the critical value for the two-tailed test of the *t*-statistics with 313 degrees of freedom and a 5% significance level). Thus, we failed to reject the null hypothesis of average producer prices being the same under these two pricing methods.

[b] The observed value of the *t*-statistic is 1.85.

the likelihood of the data being generated from each distribution. For a description of this test, see Greene (1997). To determine if the normal-jump distribution gives a better fit than a pure normal process, we test that weight discrepancy is normally distributed, and therefore there is no jump process. From equation (10), the null hypothesis of a normal distribution is written as H_0: $\mu = \sigma_J^2 = \lambda = 0$.

Nonparametric Estimation of the Probability of Regrading

The probability of regrading was also estimated using a nonparametric approach. Under the nonparametric method, the probability of regrading with and without rounding was estimated using the empirical pdf. That is, the probability of regrading is calculated by dividing the observed number of times the sum of kernels and hulls falls outside the 99–101% interval by the total number of observations in each size. The nonparametric approach is unbiased, but has a higher variance than a correctly specified parametric approach.

Empirical Results

Table 1 presents descriptive statistics of four pricing methods using paired differences on peanut producer prices. As seen from table 1, rounding introduces noise into the measurement of producer prices. The standard deviation of the difference in producer prices with and without rounding when dividing by 500g was $3.30/ton. When dividing by the actual sample weight, it was $3.35/ton.

Table 2. Parameter Estimates of the Grade Factor Linear Equations with Random Effects Models

Grade Factor / Parameter	Estimate	Std. Error	*t*-Ratio
Sound Mature Kernels (SMK):			
Intercept	−119.4897	76.77	−1.56
Sample Weight	0.8751	0.15	5.72
Weight Discrepancy	−0.7296	0.33	−2.21
Between-Groups Variance	286.3887		
Within-Groups Variance	26.0712		
Sound Splits (SS):			
Intercept	90.5113	57.73	1.57
Sample Weight	−0.1222	0.12	−1.06
Weight Discrepancy	−0.0961	0.25	−0.39
Between-Groups Variance	272.7718		
Within-Groups Variance	14.7302		
Other Kernels (OK):			
Intercept	25.7835	42.97	0.60
Sample Weight	−0.0035	0.09	−0.04
Weight Discrepancy	−0.0670	0.18	−0.37
Between-Groups Variance	54.8105		
Within-Groups Variance	8.1760		
Total Damage (TD):			
Intercept	17.5863	14.68	1.20
Sample Weight	−0.0308	0.03	−1.05
Weight Discrepancy	−0.0464	0.06	−0.78
Between-Groups Variance	1.3978		
Within-Groups Variance	0.9597		
Hulls (H):			
Intercept	−14.0701	26.74	−0.53
Sample Weight	0.2807	0.05	5.27
Weight Discrepancy	−0.0580	0.12	−0.50
Between-Groups Variance	94.5386		
Within-Groups Variance	3.1567		

Note: Expected grade factors are expressed in grams and have been calculated with no rounding since weight discrepancy is a continuous variable.

Table 2 reports the estimated parameters for the grade factor equations required for the empirical model. Note that the effect of sample weight on grade factors was only significant for sound mature kernels (SMK) and hulls (H). With little variation of the sample weight variable in the data, insignificance is expected. Weight discrepancy was found to negatively influence SMK. This finding implies most of the large errors are in measuring SMK. Table 3 details the probability of regrading with and without rounding using the empirical pdf. Rounding does tend to increase the probability of regrading as sample weights increase.

Values of the weight discrepancy ranged from −24.20 to 47.70g, with the mean and the variance being 2.77 and 4.44g, respectively. From table 4, all five parameter

Table 3. Probability of Regrading Based on the Empirical pdf

	Sample Weight			
Method of Grade Factor Determination	Size A (500g)	Size B (501.4g)	Size C (503.7g)	Size D (505g)
Rounding / 500g Sample	0.0602	0.0241	0.0241	0.0723
No Rounding / 500g Sample	0.0482	0.0120	0.0361	0.0361
Rounding / Actual Sample	0.0602	0.0482	0.1084	0.1446
No Rounding / Actual Sample	0.0602	0.0241	0.0602	0.0723

Note: The probability of regrading is calculated by dividing the number of times the sum of total kernels and hulls (grade factors measured) falls outside the 99–101% interval by the total number of observations in each size (83). Note that with 100 observations and five regrades, the 95% confidence interval would be 0.0164–0.1129.

Table 4. Parameter Estimates of the Normal-Jump Distribution of Weight Discrepancy in Grams

Parameter	Symbol	Estimate	Std. Error	*t*-Ratio
Mean of the Normal Process	α	2.2630	0.0582	38.8690
Variance of the Normal Process	σ_N^2	0.9390	0.0864	10.8640
Jump Intensity	λ	0.0705	0.0188	3.7562
Mean of the Poisson Jump	μ	5.9020	2.6662	2.2136
Variance of the Poisson Jump	σ_J^2	145.3000	58.6860	2.4760

Notes: The log of the likelihood function is −578.0905. The number of observations is 332. No rounding of grade percentages is done.

Table 5. Probability of Regrading in the Peanut Industry

Initial Sample Weight (grams)	Probability of Regrading with Rounding		Probability of Regrading without Rounding
	Theoretical Model	Empirical Model	
495	0.4900	0.5013	0.9786
496	0.3847	0.3893	0.8969
497	0.2868	0.2871	0.6193
498	0.2003	0.2136	0.2609
499	0.1422	0.1376	0.0857
500	0.0956	0.0946	0.0529
501	0.0691	0.0637	0.0499
502	0.0636	0.0524	0.0490
503	0.0653	0.0555	0.0483
504	0.0822	0.0628	0.0481
505	0.1132	0.0899	0.0563

estimates of the normal-jump distribution of weight discrepancy are statistically significant. The process associated with the sample weight lost as dust and kernels suggests graders making no errors average a loss of 2.26g of the sample. The 95% confidence interval around these estimates (0.32 to 4.20g) falls within the allowable tolerance (5g, or 1% of the sample weight). The probability of losing 5g when no large error is made was only 0.0023. This result suggests the range of the allowable tolerance is reasonable.

The estimates of the jump process imply there are about 7.05 large errors for every 100 samples, and the mean of this error is 5.90g. The probability of at least one large error is 0.0657 as calculated using the Poisson probability mass function. Figure 1 shows the normal-jump distribution of weight discrepancy given an initial sample weight of 500g.

The log-likelihood function values of the normal and Student-*t* distributions were -916.30 and $-1,119.70$. For the nested test of a normal distribution, the χ^2 test statistic of 676.4 is considerably greater than the critical value of 7.81. So, the likelihood-ratio test leads to the rejection of a normal distribution. The likelihood odds ratio[3] strongly favors the normal-jump distribution over the Student-*t* distribution.

Based on both models, the probability of regrading was calculated with and without rounding for a set of initial sample weights ranging from 495 to 505g. Both models yield very similar results. These results are reported in table 5 and are plotted in figure 2. Note, when rounding was not used, sample weights above 500g had little effect on the probability of regrading. There is little incentive to take overweight samples when rounding is not part of the grading procedure. With rounding, the probability of regrading was greatly reduced when beginning with a 502g or 503g sample. Thus, rounding does provide a strong incentive to use overweight samples. This finding offers strong evidence that the policy of rounding alone increases costs due to more frequent regrading, and reduces accuracy of grades and prices obtained with the current grading procedure.

With a 500g sample and no rounding, more than 99% of the samples which are regraded are identified as instances where the grader made a mistake. With rounding, less than 47% of the regraded samples are a consequence of grader error.[4] Without rounding, 80% of the samples on which a human error occurs are regraded, but the corresponding amount with rounding is only 63%. Therefore, regrading the wrong samples is another source of error created by rounding.

The savings from reduced regrading are modest, but they likely are enough to pay for the cost of making a change. Assuming peanut trailers average 5 tons, the 1.6 million tons produced in the United States (USDA, 2000) would provide 320,000

---

[3] With nonnested models, the negative of the log-likelihood ratio is called a likelihood odds ratio. For the Student-*t* and normal-jump models, the log-likelihood odds ratio is $1,119.70 - 578.09 = 541.61$. The interpretation is that the normal-jump model is 541.61 times more likely to originate the sample data than the Student-*t* distribution. The likelihood odds ratio is a selection criterion and not a formal nonnested test.

[4] The percentages without rounding are calculated analytically. The percentages with rounding are calculated based on the more conservative results from the empirical model. A contingency table produced from the Monte Carlo simulations is used for this purpose.
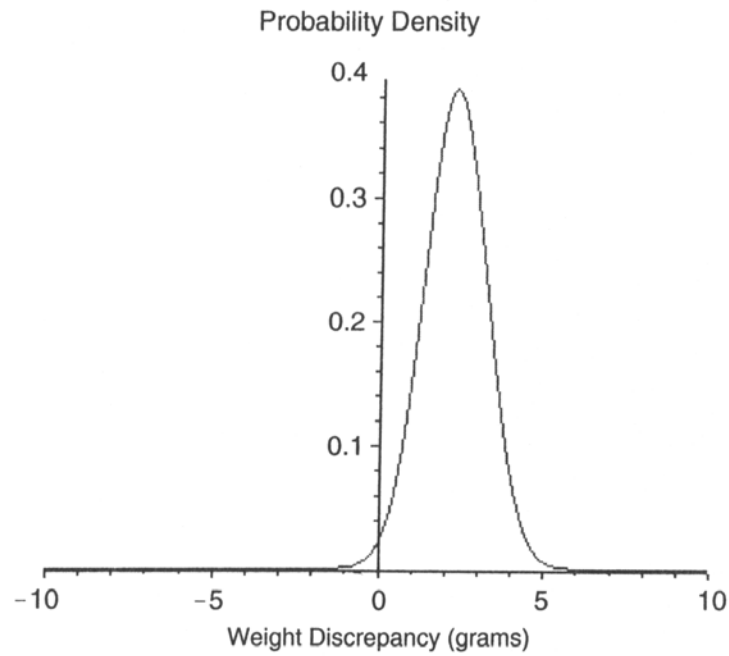
Figure 1. The normal-jump distribution of weight discrepancy in grams for a 500g sample
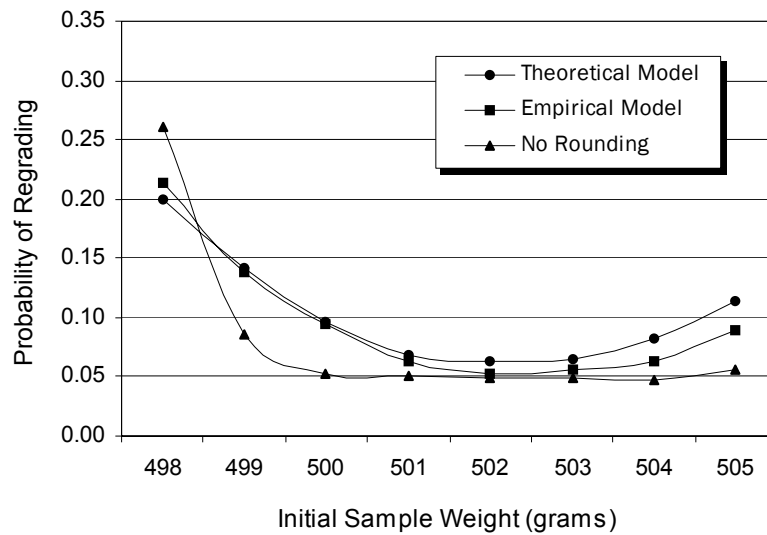

Figure 2. The probability of regrading with and without rounding

samples to be graded. If all samples were 500g, the decrease in the probability of re-grading would be 0.0417, and so there would be 13,344 fewer regrades. At $9/grade (the rate paid to private graders), this would mean reduced costs of around $120,000/year. At a 5% discount rate, the present value of an income stream of $120,000/year would be more than $2 million. Such cost savings would not be fully observed because some graders now begin with samples greater than 500g.

The incentive to take overweight samples may be the greatest concern. Biases created by taking overweight samples will not be reduced by aggregation. Graders in some grading stations never use overweight samples, so the practice of taking overweight samples moves the grading procedure away from its standards and increases the variability in grading among different buying points. This creates an inequity for producers and increases the buyer's risk.

Conclusions

The proposed normal-jump distribution appears to be a good model for a process composed of many small errors and occasional large errors. The current practice of rounding of grade factors to the nearest whole number introduces noise, increases costs due to more frequent regrading, causes the wrong samples to be regraded, and provides a major incentive for graders to take overweight samples. A very low-cost way to improve the peanut grading procedure would be to discontinue rounding. Stopping the rounding of grade factors could pay for itself since it would reduce the frequency of regrading. In this technology era of sophisticated computers and calculators, the practice of rounding to whole numbers is certainly no longer necessary.

References

Akgiray, V., and G. Booth. (1986, May). "Stock price processes with discontinuous time paths: An empirical examination." *Financial Review* 21, 163–184.

———. (1988). "Mixed diffusion-jump process modeling of exchange rate movements." *Review of Economics and Statistics* 70, 631–637.

Anderson, K. B. (1998, November). Professor, Department of Agricultural Economics, Oklahoma State University. Personal communication regarding accuracy in peanut grading procedure.

Char, B. W., K. O. Geddes, G. H. Gonnet, B. L. Leong, M. B. Monagan, and S. M. Watt. (1993). *First Leaves: A Tutorial Introduction to Maple V*. New York: Springer-Verlag.

Dowell, F. E., C. R. Meyer, and R. P. Konstance. (1994). "Identification of feasible grading system improvements using systems engineering." *Applied Engineering in Agriculture* 10, 717–723.

Greene, W. *Econometric Analysis*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1997.

Oklahoma Federal/State Inspection Service. (1995, August 14). *Standard Operating Procedures for Peanut Grading Operations*. Pub. No. SOP-246, Agricultural Products Division, Fruit and Vegetable Section, Department of Agriculture, Oklahoma City.

Oldfield, G. S., R. J. Rogalski, and R. A. Jarrow. (1977). "An autoregressive jump process for common stock returns." *Journal of Financial Economics* 5, 389–418.

Powell, J. H., Jr., H. T. Sheppard, and F. E. Dowell. (1994, December). "An automated data collection system for use in grading of farmers' stock peanuts." Paper presented at the 1994 International Winter Meeting sponsored by the American Society of Agricultural Engineers (ASAE), Atlanta, Georgia, 13–16 December 1994.

SAS Institute, Inc. *Getting Started with PROC MIXED.* Cary, NC: SAS Institute, Inc., 1994.

Steigert, K. W., and B. W. Brorsen. (1996). "The distribution of futures prices: Diffusion-jump vs. generalized beta-2." *Applied Economics Letters* 3, 303–305.

Tsai, Y. J., F. E. Dowell, J. I. Davidson, Jr., and R. J. Cole. (1993, December). "Accuracy and variability in sampling and grading Florunner farmers' stock peanuts." *Oléagineux* 48, 511–514.

Tucker, A. L., and L. Pond. (1988, November). "The probability distribution of foreign exchange price changes: Tests of candidate processes." *Review of Economics and Statistics* 70, 638–647.

U.S. Department of Agriculture, Federal/State Inspection Service. (1996, August). *Farmers' Stock Peanuts Inspection Instructions Handbook*, Update 133. USDA/FSIS, Agricultural Marketing Service, Fruit and Vegetable Division, Fresh Products Branch, Washington, DC.

U.S. Department of Agriculture, National Agricultural Statistics Service. (2000). *Agricultural Statistics 1999.* Washington, DC: U.S. Government Printing Office.

Whitaker, T. B., J. W. Dickens, and F. G. Giesbretch. (1991). "Variability associated with determining grade factors and support price of farmers' stock peanuts." *Peanut Science* 18, 122–126.

Whitaker, T. B., J. Wu, F. E. Dowell, W. M. Hagler, Jr., and F. G. Giesbrecht. (1994). "Effects of sample size and sample acceptance level on the number of aflatoxin-contaminated farmers' stock lots accepted and rejected at the buying point." *Journal of the Association of Official Analytical Chemists* 77, 1672–1680.

White, K. J. (1993). *SHAZAM User's Reference Manual, Version 7.0*, 2nd printing. Vancouver, Canada: McGraw-Hill Book Co.

Widrow, B., I. Kollár, and M.-C. Liu. (1996, April). "Statistical theory of quantization." *IEEE Transactions on Instrumentation and Measurement* 45, 353–361.