Effect of Sample Size and Method of Sampling Pig Weights on the Accuracy and Precision of Estimating the Distribution of Pig Weights in a Population[1,2]

C. B. Paulk, G. L. Highland[3], M. D. Tokach, J. L. Nelssen, S. S. Dritz[4], R. D. Goodband, J. M. DeRouchey, and K. D. Haydon[5]

Summary

Producers have adopted marketing strategies such as topping to help reduce economic losses from weight discounts, but they are still missing target weights and incurring discounts. We have previously determined the accuracy of sampling methods producers use to estimate the mean weight of the population. Although knowing the mean weight is important, understanding how much variation or dispersion exists in individual pig weights within a group can also enhance a producer's ability to determine the optimal time to top pigs. In statistics and probability theory, the amount of variation in a population is represented by the standard deviation; therefore, our objective is to determine the sample size and method that is optimal for estimating the standard deviation of BW for a group of pigs in a barn.

Using a computer program developed in R (R Foundation for Statistical Computing, Vienna, Austria), we were able to generate 10,000 sample standard deviations for different sampling procedures on 3 different datasets. Using this program, we evaluated weighing: (1) a completely random sample of 10 to 200 pigs from the barn, (2) an increasing number of pigs per pen from 1 to 15 pigs and increasing the number of pens until all pens in the barn had been sampled, and (3) selecting the heaviest and lightest pig (determined visually) in each pen, and subtracting the lightest weight from the heaviest weight and dividing by 6. For all 3 datasets, increasing the sample size of a completely random sample from 10 to 200 pigs decreased the range between the upper and lower confidence intervals (CI) when estimating the standard deviation; however, this occurred at a diminishing rate. For the barn with the most variation, increasing the number of pens sampled while keeping constant the total number of pigs sampled led to a reduction in range between the upper and lower CI by 7, 6, and 31% for Datasets A, B, and C, respectively. Sampling method 3 resulted in a reduction of the range between the upper and lower CI from 9 to 62% for the 3 datasets. These data indicated that the distribution of pig weights can be practically estimated by weighing the heaviest and lightest pigs in 15 pens.

Key words: finishing pig, standard deviation estimation, sample size

Introduction

Despite adopting marketing strategies such as topping to help reduce economic losses at the processing plant, swine producers are often missing target weights and incurring substantial weight discounts. We have previously determined the accuracy of sampling methods producers use to estimate the mean weight of the population. Although knowing the mean weight is important, understanding how much variation or dispersion exists in individual pig weights from the mean weight can also enhance a producer's ability to maximize economic return when marketing pigs. Knowing the distribution allows producers to better estimate the ideal timing for removing pigs from a barn. In statistics and probability theory, the amount of variation in a population is represented by the standard deviation; therefore, our objective was to determine the optimal sample size and method for estimating the standard deviation of weights for the population of pigs in the barn.

Procedures

A total of 3 datasets (A, B, and C) in which all pigs in the barn had been weighed individually were used to evaluate sample size and method of sampling on the precision of estimating the variation in pig weights in the barn. The first method of sampling tested was a completely random sample of the barn that disregarded pen arrangements. Samples of different sizes were taken (10, 20, 30 pigs, etc.). The second sampling method tested increasing the number of pigs sampled per pen from 1 to 15 pigs, then increasing the number of pens until all pens had been sampled. The third sampling method consisted of selecting the heaviest and lightest pig (determined visually) from 15 pens (30 pigs total) and dividing the difference in weight between the lightest and heaviest pigs in the total sample by 6.

Dataset A was derived from Groesbeck et al. (2007[6]). Dataset A (Figure 1) contained a total of 1,260 pigs from 48 pens with 23 to 28 pigs per pen. The mean, median, standard deviation and CV of the population were 253.0 lb, 254 lb, 32.8 lb, and 13.0%, respectively. Datasets B and C were obtained for the purposes of this experiment. Dataset B was obtained from a commercial site in northern Iowa. The finishing facility utilized PIC C42 × PIC 359 pigs that were classified as healthy by the farm veterinarian. The barn was filled with pigs over a 1-wk period, and pigs were gate cut as they came off the truck to randomly place them in pens. For dataset B (Figure 2), a total of 1,261 pigs were weighed (population mean = 213.5 lb, median = 214 lb, standard deviation = 21.5 lb, and CV = 10.1%) from 19 pens with 56 to 81 pigs per pen. The 20th pen was used as a recovery pen and was not used for analysis. Dataset C was derived from a different commercial site in northern Iowa that consisted of pigs (Genetiporc F25 × G performer boar) weaned during a porcine reproductive and respiratory syndrome (PRRS) outbreak at the sow farm. The barn was filled with pigs over a 1-wk period, and pigs were gate cut as they came off the truck into pens. For Dataset C (Figure 3), a total of 1,069 pigs were weighed (population mean = 222.4 lb, median = 224 lb, standard deviation = 32.0 lb, and CV = 14.4%) from 40 pens with 20 to 35 pigs per pen. The barn did not have a recovery pen for sick pigs; therefore, all pens were used for analysis.

A program was coded using R (R Foundation for Statistical Computing, Vienna, Austria) to demonstrate the error that varying sample sizes and methods of selecting pig weights have on the

[6] Groesbeck, G. N., G. Armbuster, M. D. Tokach, R. D. Goodband, J. M. DeRouchey, and J. L. Nelssen. 2007. Influence of Pulmotil, Tylan, and Paylean on pig growth performance and weight variation. American Association of Swine Veterinarians Proceedings, pp. 235–238.

estimation of the standard deviation of a population. For the first method of sampling, the program was designed to take a completely random sample of the designated sample size, disregarding pen arrangements, and calculate the standard deviation of this sample. The standard deviation was calculated as:

Standard deviation = $\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2}$, where n is the sample size, $\{x_1, x_2, \ldots x_n\}$ are the observed values of the sample items, and $\bar{x}$ is the mean value of these observations.

The program conducted the sampling technique 10,000 times, generating 10,000 sample standard deviation calculations for each sample size (10, 20, 30 pigs, etc.) by randomly selecting the desired number of pig weights from the population. The 10,000 sample standard deviations for each sample size were sorted from least to greatest. A 95% confidence interval (CI) was generated by selecting the $9{,}751^{st}$ observation (upper CI) and the 250th observation (lower CI). The distances between the upper and lower CIs represent the range of the mean estimations. A similar analysis was conducted using R for the second method, but the second sampling method tested the sampling error among a varying number of pigs within varying numbers of pens, with 1 to 15 pigs sampled from 1 to all of the pens.

A similar analysis was conducted using R to determine the error associated with sampling method 3. Personnel trained in selecting pigs (marketers) provided by Suidae Health and Production, Algona, IA, chose the heaviest and lightest pigs in each pen. One marketer, marketer 1, was provided for Dataset B, and two marketers, marketers 2 and 3, were provided for Dataset C. Selection accuracy was incorporated into sampling method 3 for Dataset A based on the selection accuracy of the 2 marketers from Dataset C. The probability for selecting the 1st, 2nd, 3rd, 4th, or 5th heaviest pig was 50, 25, 15, 5, and 5%, respectively, and the probability for selecting the 1st, 2nd, 3rd, 4th, or 5th lightest pig was 70, 15, 5, 5, and 5%, respectively. These were chosen because Datasets A and C had similar pen arrangements. To account for selection accuracy in the simulations, a rank was assigned to the heaviest and lightest pig selected by the marketer in each pen. For each pen selected, a rank was randomly selected; therefore, for Dataset A, if the 1st pen randomly selected was pen 8, one pig selected from pen 8 would have a 50, 25, 15, 5, and 5% chance of being either the 1st, 2nd, 3rd, 4th, or 5th heaviest pig and the other pig selected would have a 70, 15, 5, 5, and 5% chance of being either the 1st, 2nd, 3rd, 4th, or 5th lightest pig, respectively.

Results and Discussion
Notably, the random samples were generated using a computer program, but those samples taken from the barn are not truly random unless pigs are individually identified and preselected, rather than being selected by the marketer.

For all 3 datasets, increasing the sample size of a completely random sample from 10 to 200 pigs decreased the range between the upper and lower CI when estimating the standard deviation (Figures 5, 6, and 7). A majority of the improvement in the precision of the estimation occurred when the sample size increased from 10 to 90 pigs (Table 1). The difference in accuracy of sample size between the different datasets is also important to note. This could result from the difference in the variation of each dataset (Figures 1, 2, and 3); for example, Dataset B had less variation, so fewer pigs needed to be sampled to achieve a similar CI range.

Individual pen means ranged from 253 to 276, 186 to 222, and 180 to 228 lb for Datasets A, B, and C, respectively. Individual pen standard deviations ranged from 19 to 47, 15 to 25, and 16 to 44 lb for Datasets A, B, and C, respectively. As both the number of pigs and pens were increased when sampling, the range or distance between the upper and lower CI decreased (Figures 8, 9, 10 and Tables 2, 3, and 4). Increasing the number of pens sampled while keeping the total number of pigs sampled constant at 30 pigs led to a reduction in range between the upper and lower CI (Table 5). For Datasets A and C, when sampling 15 pigs from 2 pens, the estimated range between the upper and lower CI was 19.9 and 25.2 lb, respectively; however, when sampling 1 pig from 30 pens, the range between the upper and lower CI was 18.5 and 17.5 lb for Datasets A and C, respectively. For Dataset B, when sampling 15 pigs from 2 pens, the estimated range between the upper and lower CI was 12.1 lb, but when sampling 1 pig from 30 pens, the range between the upper and lower CI was 11.4 lb. Therefore, increasing the number of pens used when sampling the barn can improve the range between the upper and lower CI by 7, 6, and 31% for Datasets A, B, and C, respectively, but a major improvement occurred only in Dataset C because Dataset C had a larger difference between individual pen means and standard deviations. Because the distribution of pig weights across pens is not known, taking a random sample from an increasing number of pens is recommended when estimating the distribution of pig weights in the barn.

When asked to identify the heaviest pig in the pen, marketers 1, 2, and 3 identified the heaviest pig in 47.4, 43.5, and 55.0% of the pens and the 2nd heaviest pig in 5.3, 35.0, and 25.0% of the pens, respectively (Table 1). The pigs identified by marketers 1, 2, and 3 were within the actual 5 heaviest pigs in 68, 100, and 95% of the pens, respectively. When asked to select the lightest pig, marketers 1, 2, and 3 identified the lightest pig in 57.9, 75.0, and 68.4% of the pens and the 2nd lightest pig in 21.1, 17.5, and 10.5% of the pens, respectively (Table 6). The pigs identified by marketers 1, 2, and 3 were within the actual 5 lightest pigs in 89.5, 100, and 100% of the pens, respectively.

Selecting the heaviest and lightest pigs in 15 pens and dividing the difference between the heaviest and lightest pig of the 30 selected pigs by 6 resulted in a reduction of the range between the upper and lower CI. Amongst the various datasets, the range was reduced from 9 to 62% compared with randomly selecting 2 pigs from 15 pens. Sampling method 3 is expected to be a good estimator of the standard deviation, because in a population that approximates a normal distribution, 99.9% of observations are should be within plus or minus 3 standard deviations of the mean, a total of 6 standard deviations between the heaviest and lightest observation; consequently, selecting the heaviest and lightest weight of the distribution and dividing by 6 should approximate the standard deviation of the population.

Sample size, method, variation, and distribution of pigs within a barn can substantially affect the precision of estimating the distribution of pig weights. As expected, sample size to obtain similar CI estimates is reduced if the population is less variable. Finally, these data indicate that the distribution of pig weights can be estimated practically by weighing the heaviest and lightest pigs in 15 pens.

Figure 1. Histogram of Dataset A, a total of 1,260 pigs (mean = 253.0 lb, median = 254 lb, standard deviation = 32.8 lb, and CV = 12.98%) with 23 to 28 pigs per pen and a total of 48 pens.
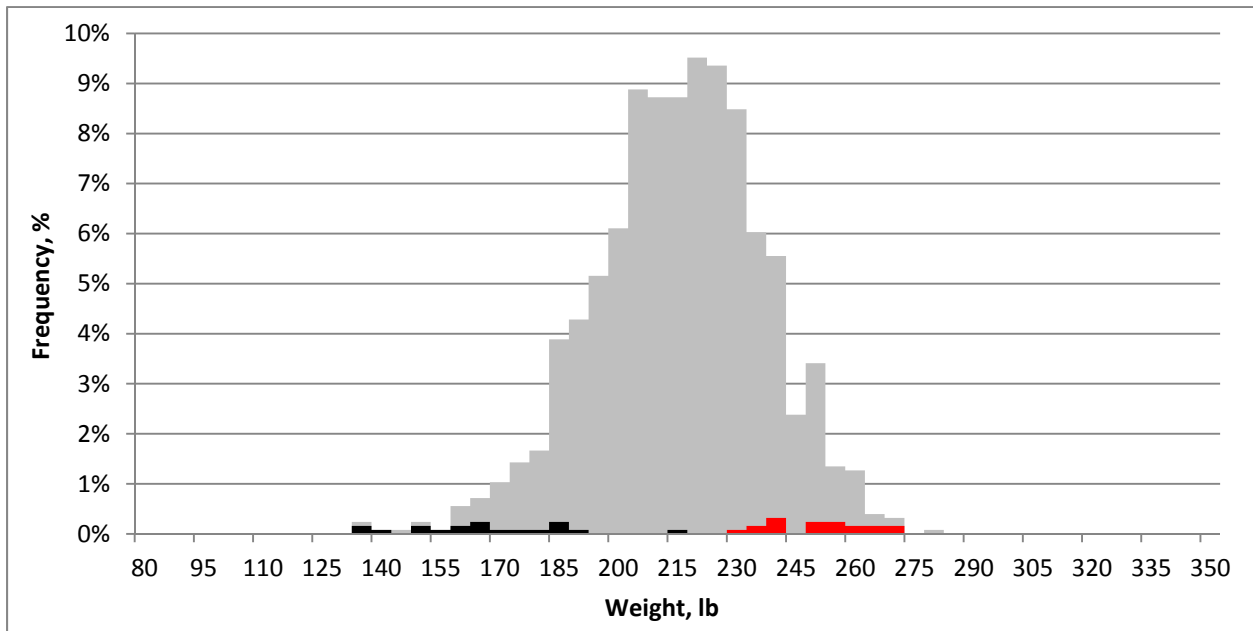
Figure 2. Histogram of Dataset B and marketer 1's selections. A total of 1,261 pigs were weighed (population mean = 213.5 lb, median = 214 lb, standard deviation = 21.5 lb, and CV = 10.1%), with 19 pens and 56 to 81 pigs per pen. The marketer selected the heaviest and lightest pig in each pen. The 2 histograms of the marketer's selections are imposed on top of the population histogram.
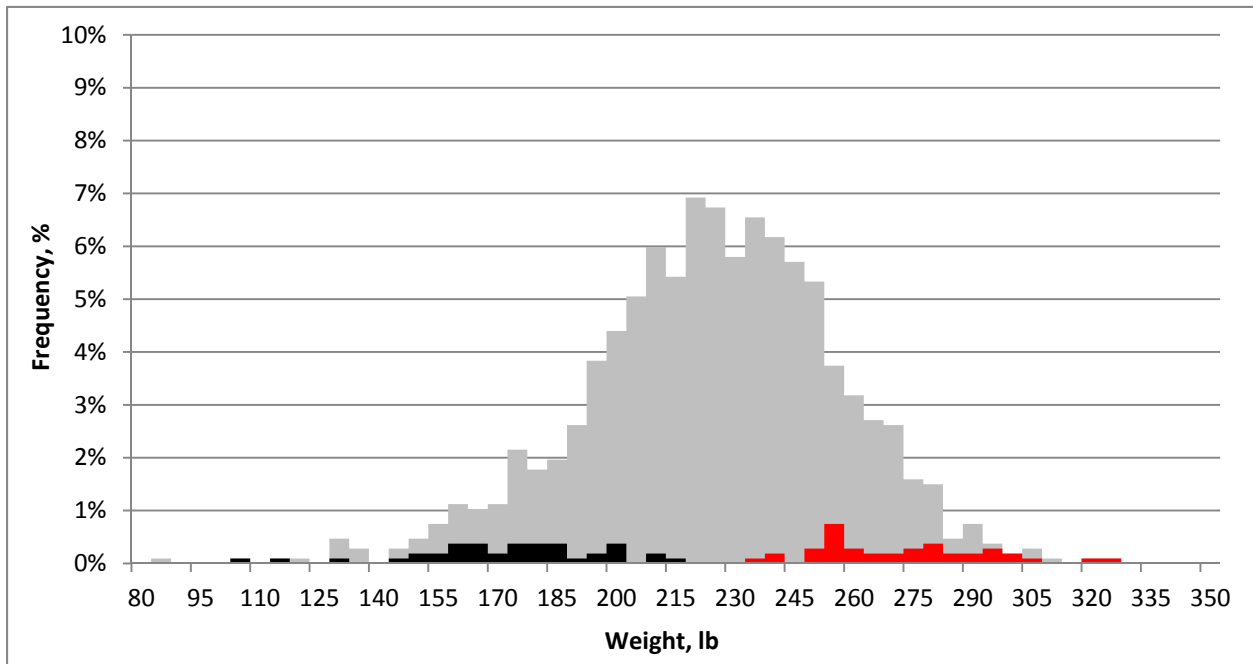
Figure 3. Histogram of Dataset C and marketer 2's selections. A total of 1,069 pigs were weighed (population mean = 222.4 lb, median = 224 lb, standard deviation = 32.0 lb, and CV = 14.4%), with 40 pens and 20 to 35 pigs per pen. The marketer selected the heaviest and lightest pig in each pen. The histograms of the lightest and heaviest of the selections are imposed on top of the population histogram.
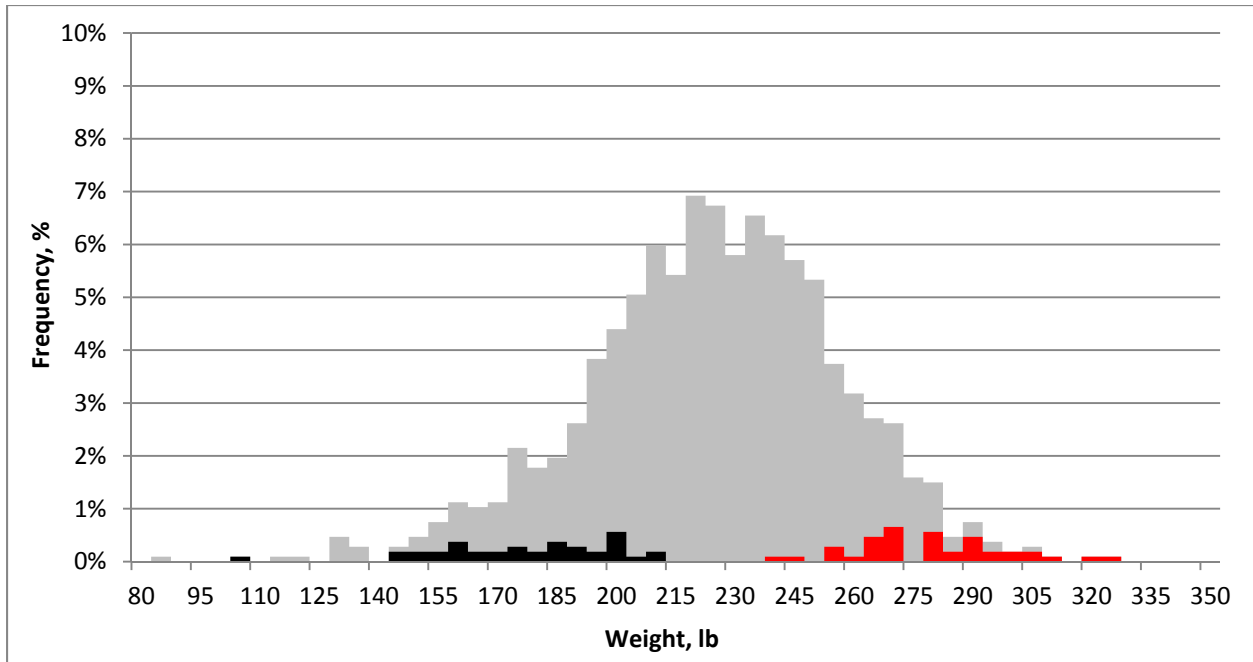
Figure 4. Histogram of Dataset C and marketer 3's selections. A total of 1,069 pigs were weighed (population mean = 222.4 lb, median = 224 lb, standard deviation = 32.0 lb, and CV = 14.4%), with 40 pens and 20 to 35 pigs per pen. The marketer selected the heaviest and lightest pig in each pen. The histograms of the lightest and heaviest selections are imposed on top of the population histogram.

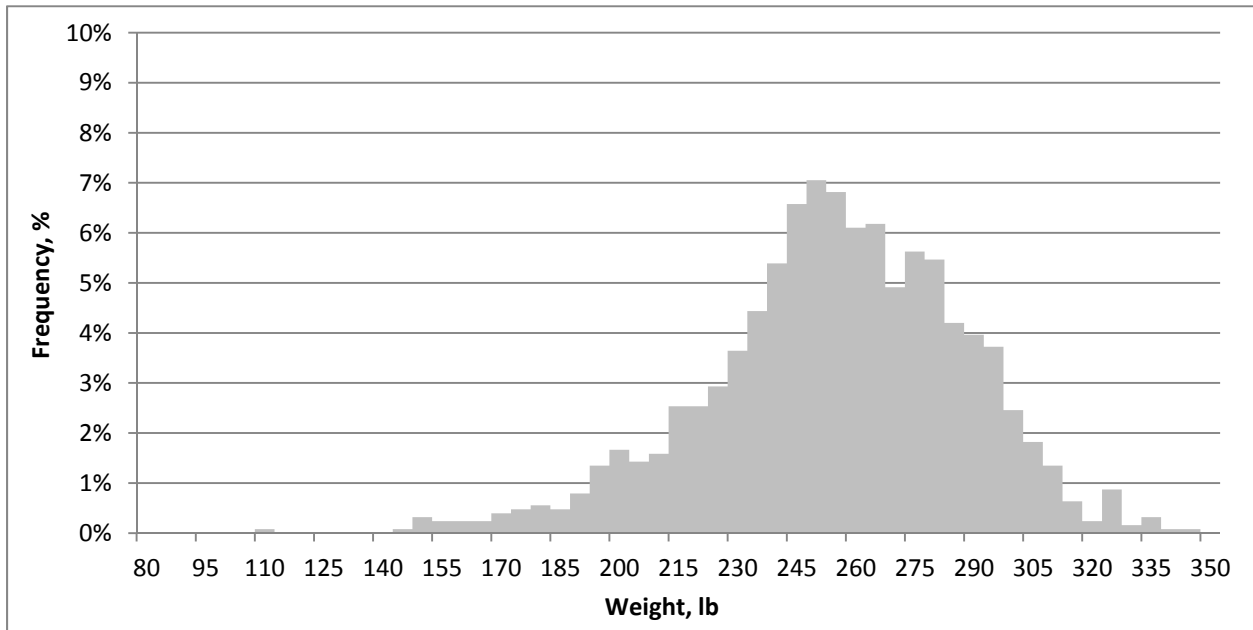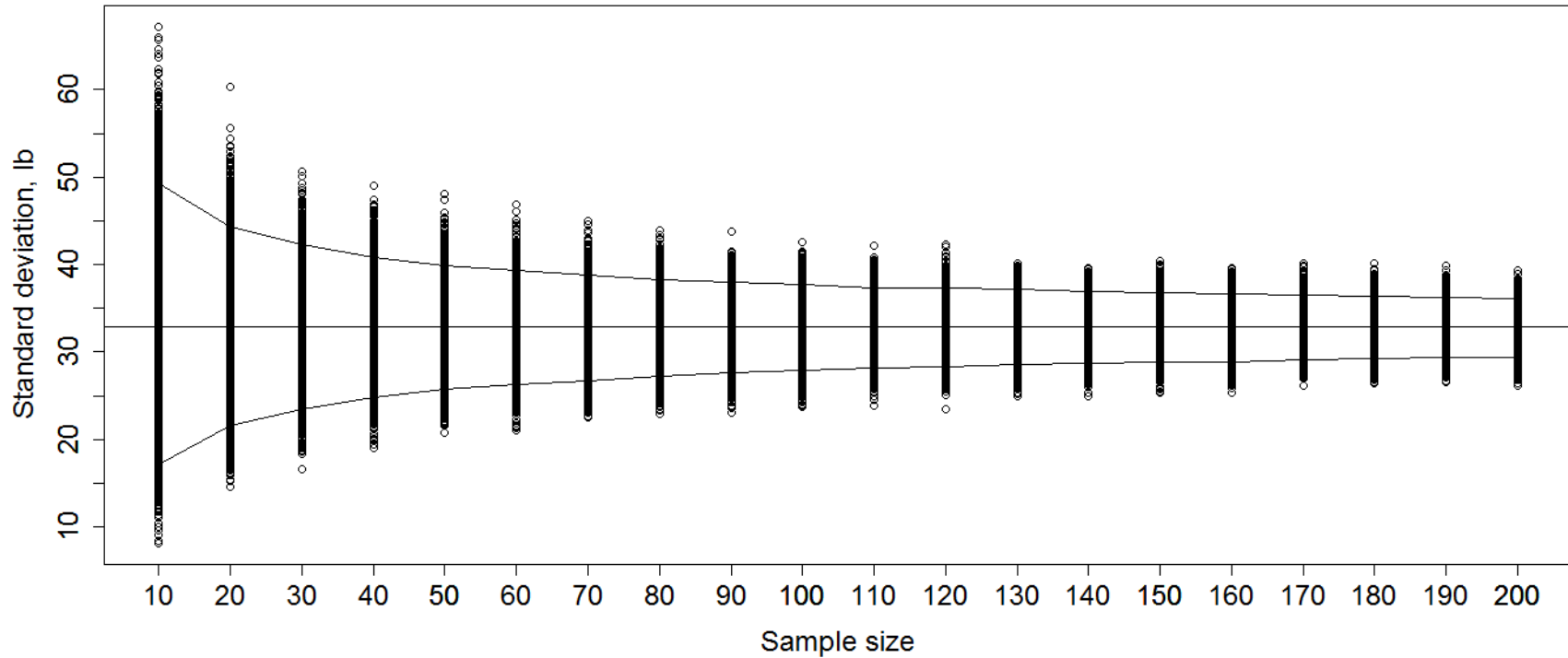**Figure 5.** For dataset A, individual pig weights were collected on a total of 1,260 pigs (mean = 253.0 lb, median = 254 lb, standard deviation = 32.8 lb, and CV = 12.98%) with 23 to 28 pigs per pen. The datasets were then analyzed by taking random samples, disregarding pen arrangements, of different sample sizes (10, 20, 30, etc.) and calculating the standard deviation. This operation was completed 10,000 times for each sample size. Each point represents the standard deviation calculated for the respective sample. Reference lines representing the 95% confidence interval have been drawn, and the center line represents the actual population standard deviation.
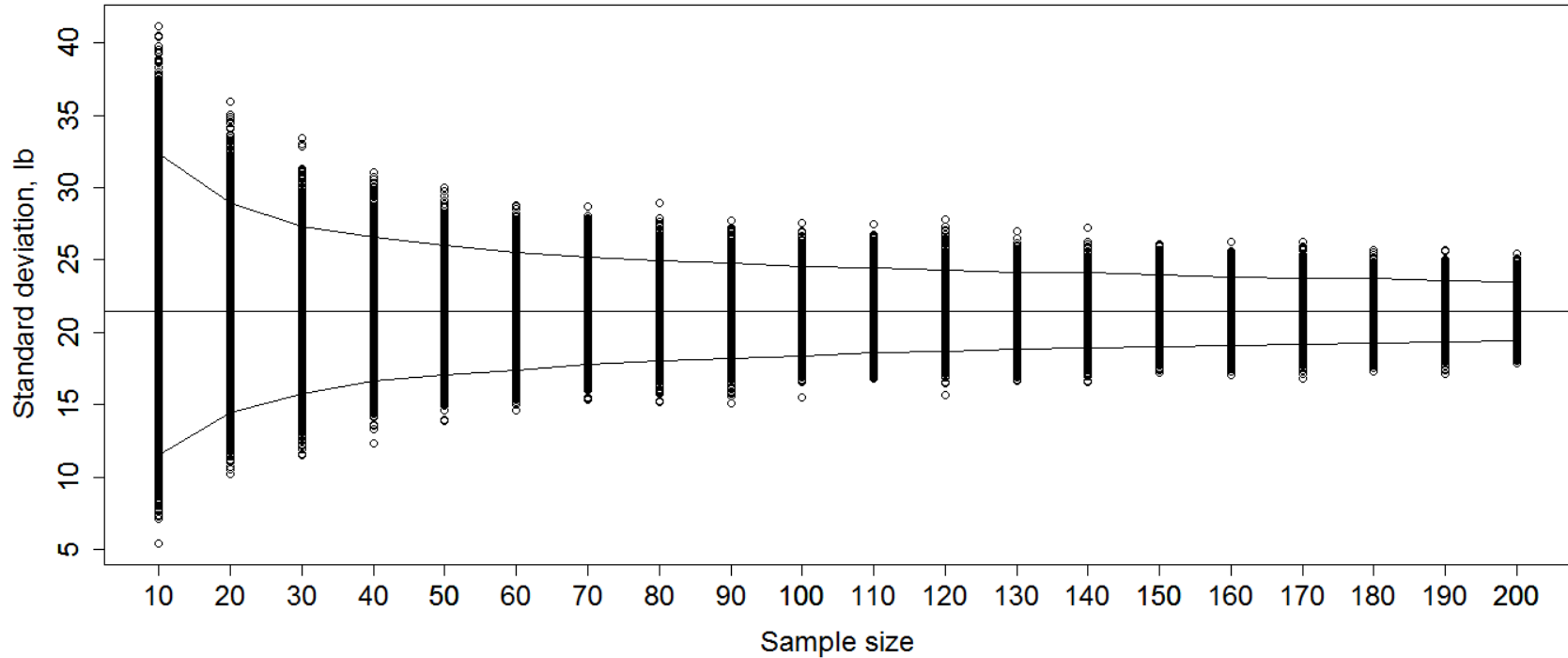
Figure 6. For Dataset B, individual pig weights were collected on a total of 1,261 pigs (population mean = 213.5 lb, median = 214 lb, standard deviation = 21.5 lb, and CV = 10.1%) from 19 pens with 56 to 81 pigs per pen. The datasets were then analyzed by taking random samples, disregarding pen arrangements, of different sample size (10, 20, 30, etc.) and calculating the standard deviation. This operation was completed 10,000 times for each sample size. Each point represents the standard deviation calculated for the respective sample. Reference lines representing the 95% confidence interval have been drawn, and the center line represents the actual population standard deviation.
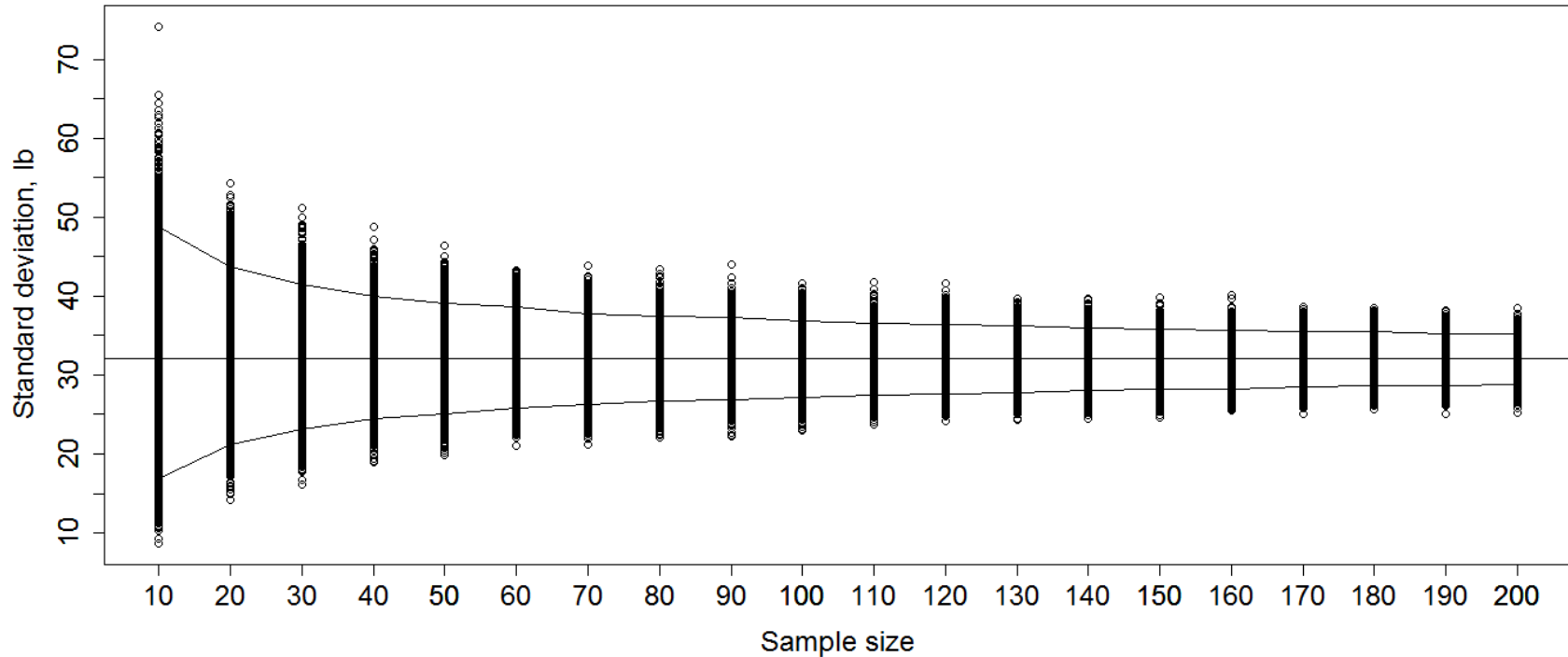
Figure 7. For Dataset C, individual pig weights were collected on a total of 1,069 pigs weighed (population mean = 222.4 lb, median = 224 lb, standard deviation = 32.0 lb, and CV = 14.4%) from 40 pens with 20 to 35 pigs per pen. The datasets were then analyzed by taking random samples, disregarding pen arrangements, of different sample size (10, 20, 30, etc.) and calculating the standard deviation. This operation was completed 10,000 times for each sample size. Each point represents the standard deviation calculated for the respective sample. Reference lines representing the 95% confidence interval have been drawn, and the center line represents the actual population standard deviation.

Table 1. The mean standard deviation, upper confidence interval (CI), lower confidence interval, and range of estimates of the standard deviation when taking a completely random sample of 30, 60, 90, or 120 pigs from the datasets

Sampling method	Mean of 10,000 simulations[1]	Upper CI	Lower CI	Range
Dataset A[2]				
30 pigs	32.5	42.2	23.5	18.7
60 pigs	32.6	39.3	26.4	13.0
90 pigs	32.7	38.0	27.6	10.4
120 pigs	32.8	37.3	28.3	9.0
Dataset B[3]				
30 pigs	21.3	27.3	15.7	11.6
60 pigs	21.4	25.5	17.4	8.2
90 pigs	21.4	24.8	18.2	6.5
120 pigs	21.5	24.3	18.7	5.7
Dataset C[4]				
30 pigs	31.7	41.4	23.2	18.2
60 pigs	31.9	38.6	25.9	12.7
90 pigs	32.0	37.3	26.9	10.4
120 pigs	32.0	36.4	27.7	8.8

[1] The standard deviation was calculated for each of the generated samples, and the mean of the 10,000 generated standard deviation estimates was determined.

[2] A total of 1,260 pigs (mean = 253.0 lb, median = 254 lb, standard deviation = 32.8 lb, and CV = 12.98%) with 23 to 28 pigs per pen and a total of 48 pens.

[3] A total of 1,261 pigs (population mean = 213.5 lb, median = 214 lb, standard deviation = 21.5 lb, and CV = 10.1%) with 56 to 81 pigs per pen and atotal of 19 pens.

[4] A total of 1,069 pigs weighed (population mean = 222.4 lb, median = 224 lb, standard deviation = 32.0 lb, and CV = 14.4%) with 40 pens and 20 to 35 pigs per pen.
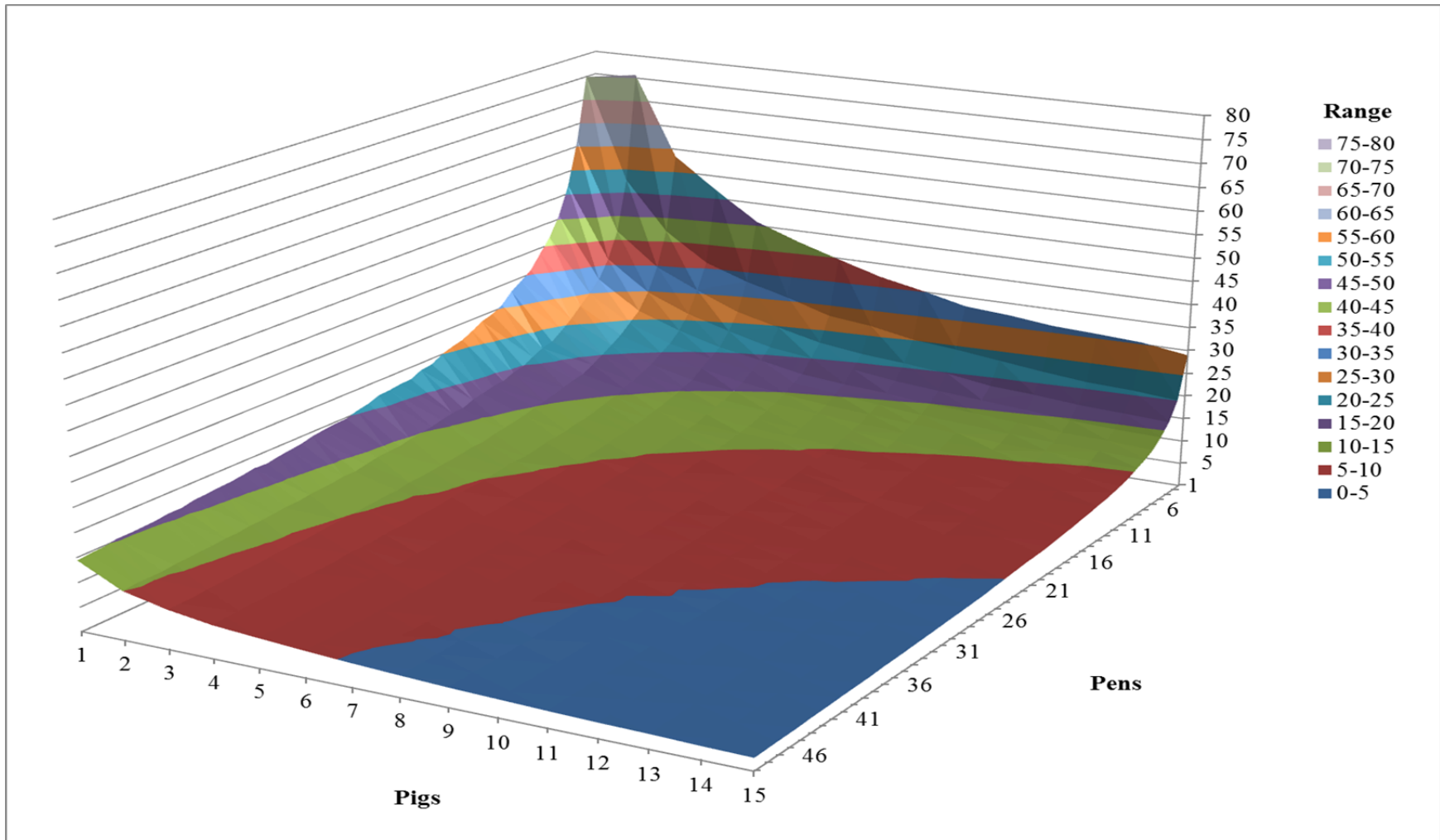
Figure 8. For Dataset A, individual pig weights were collected on a total of 1,260 pigs (actual population weight = 253.0 lb and CV = 12.98%) from 48 pens with 23 to 28 pigs per pen. The dataset was analyzed by estimating the overall standard deviation by altering the number of pigs selected within pens, and total number of pens sampled. This operation was completed 10,000 times for each sampling method, and the range or difference between the upper and lower CI was calculated. Each point on this graph shows the range between the upper and lower CI, represented in pounds.

Table 2. The range between the upper and lower confidence interval (CI) for varying pigs and pen as presented in Figure 7 (Dataset A)[1]

Pens, n	Number of pigs from each pen														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		76	58	52	46	43	40	37	35	33	32	31	31	30	29
2	75	53	42	36	33	31	28	27	25	24	23	22	21	21	20
3	61	42	35	31	27	25	23	21	21	20	18	18	17	17	16
4	52	36	30	26	23	21	20	19	18	17	16	16	15	14	14
5	46	32	26	23	21	19	17	17	16	15	14	14	13	13	12
6	42	30	25	21	19	18	16	15	14	14	13	12	12	12	11
7	39	28	23	20	18	16	15	14	13	13	12	11	11	10	10
8	36	26	21	18	16	15	14	13	12	12	11	11	10	10	9
9	34	24	20	17	15	14	13	12	12	11	10	10	10	9	9
10	33	23	19	16	15	13	12	11	11	10	10	9	9	9	8
11	31	22	18	16	14	13	12	11	10	10	9	9	8	8	8
12	29	21	17	15	13	12	11	11	10	9	9	8	8	8	8
13	29	20	17	14	13	11	11	10	9	9	8	8	8	7	7
14	27	19	16	14	12	11	10	10	9	9	8	8	7	7	7
15	26	18	15	13	12	11	10	9	9	8	8	8	7	7	7
16	26	18	15	13	11	10	9	9	8	8	7	7	7	7	6
17	25	18	14	12	11	10	9	9	8	8	7	7	7	6	6
18	24	17	14	12	11	10	9	8	8	7	7	7	6	6	6
19	23	16	13	12	10	9	9	8	8	7	7	6	6	6	6
20	22	16	13	11	10	9	8	8	7	7	6	6	6	6	6
21	22	16	13	11	10	9	8	8	7	7	6	6	6	6	5
22	22	15	12	11	9	9	8	7	7	7	6	6	6	5	5
23	22	15	12	11	9	8	8	7	7	6	6	6	5	5	5
24	21	15	12	10	9	8	8	7	7	6	6	6	5	5	5
25	20	14	12	10	9	8	7	7	6	6	6	5	5	5	5
26	20	14	11	10	9	8	7	7	6	6	6	5	5	5	4
27	20	14	11	10	8	8	7	6	6	6	5	5	5	5	4
28	19	14	11	9	8	7	7	6	6	6	5	5	5	4	4
29	19	13	11	9	8	7	7	6	6	5	5	5	5	4	4
30	18	13	11	9	8	7	6	6	6	5	5	5	4	4	4
31	18	13	10	9	8	7	6	6	5	5	5	5	4	4	4
32	18	13	10	9	8	7	6	6	5	5	5	4	4	4	4
33	17	12	10	9	8	7	6	6	5	5	5	4	4	4	4
34	18	12	10	8	7	7	6	6	5	5	5	4	4	4	4
35	17	12	10	8	7	7	6	5	5	5	4	4	4	4	4
36	17	12	9	8	7	6	6	5	5	5	4	4	4	4	3
37	17	12	9	8	7	6	6	5	5	5	4	4	4	4	3
38	16	12	9	8	7	6	6	5	5	4	4	4	4	3	3

39	16	11	9	8	7	6	6	5	5	4	4	4	4	3	3
40	16	11	9	8	7	6	5	5	5	4	4	4	3	3	3
41	16	11	9	8	7	6	5	5	5	4	4	4	3	3	3
42	16	11	9	7	7	6	5	5	4	4	4	4	3	3	3
43	15	11	9	7	6	6	5	5	4	4	4	3	3	3	3
44	15	11	8	7	6	6	5	5	4	4	4	3	3	3	3
45	15	10	8	7	6	6	5	5	4	4	4	3	3	3	3
46	15	10	8	7	6	5	5	4	4	4	4	3	3	3	3
47	15	10	8	7	6	5	5	4	4	4	3	3	3	3	3
48	14	10	8	7	6	5	5	4	4	4	3	3	3	3	2

[1] Colors match the color scheme in Figure 8, representing a range of 5 lb for each color.

Figure 9. For Dataset B, individual pig weights were collected on a total of 1,261 pigs (population mean = 213.5 lb, median = 214 lb, standard deviation = 21.5 lb, and CV = 10.1%) from 19 pens with 56 to 81 pigs per pen. The dataset was analyzed by altering the number of pigs selected within pens, and total number of pens sampled. This operation was completed 10,000 times for each sampling method, and the range or difference between the upper and lower CI was calculated. Each point on this graph shows the range between the upper and lower CI, represented in pounds.

Table 3. The range between the upper and lower confidence interval (CI) for varying pigs and pen as presented in Figure 7 (Dataset B)[1]

Pens, n	Number of pigs from each pen														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		45	37	32	28	26	24	23	21	20	19	18	18	17	16
2	49	33	27	24	21	19	18	17	16	15	15	14	14	13	13
3	39	27	22	19	17	16	15	14	13	13	13	12	12	11	11
4	33	23	19	17	15	14	13	12	12	11	11	11	10	10	10
5	29	21	17	15	13	12	12	11	10	10	10	9	9	9	9
6	27	19	15	14	12	11	11	10	9	9	9	8	8	8	8
7	25	17	14	13	11	10	10	9	9	8	8	8	8	7	7
8	23	16	13	12	11	10	9	8	8	8	7	7	7	7	7
9	22	15	12	11	10	9	8	8	8	7	7	7	6	6	6
10	20	14	12	10	9	8	8	8	7	7	6	6	6	6	6
11	20	14	11	10	9	8	8	7	7	6	6	6	6	6	5
12	19	13	11	9	8	8	7	7	6	6	6	5	5	5	5
13	18	12	10	9	8	7	7	6	6	6	6	5	5	5	5
14	17	12	10	8	7	7	6	6	6	5	5	5	5	5	4
15	16	11	9	8	7	7	6	6	5	5	5	5	4	4	4
16	16	11	9	8	7	6	6	5	5	5	5	4	4	4	4
17	15	11	9	7	7	6	6	5	5	5	4	4	4	4	4
18	15	10	8	7	6	6	5	5	5	4	4	4	4	4	4
19	14	10	8	7	6	6	5	5	4	4	4	4	4	3	3

[1] Colors match the color scheme in Figure 9, representing a range of 5 lb for each color.
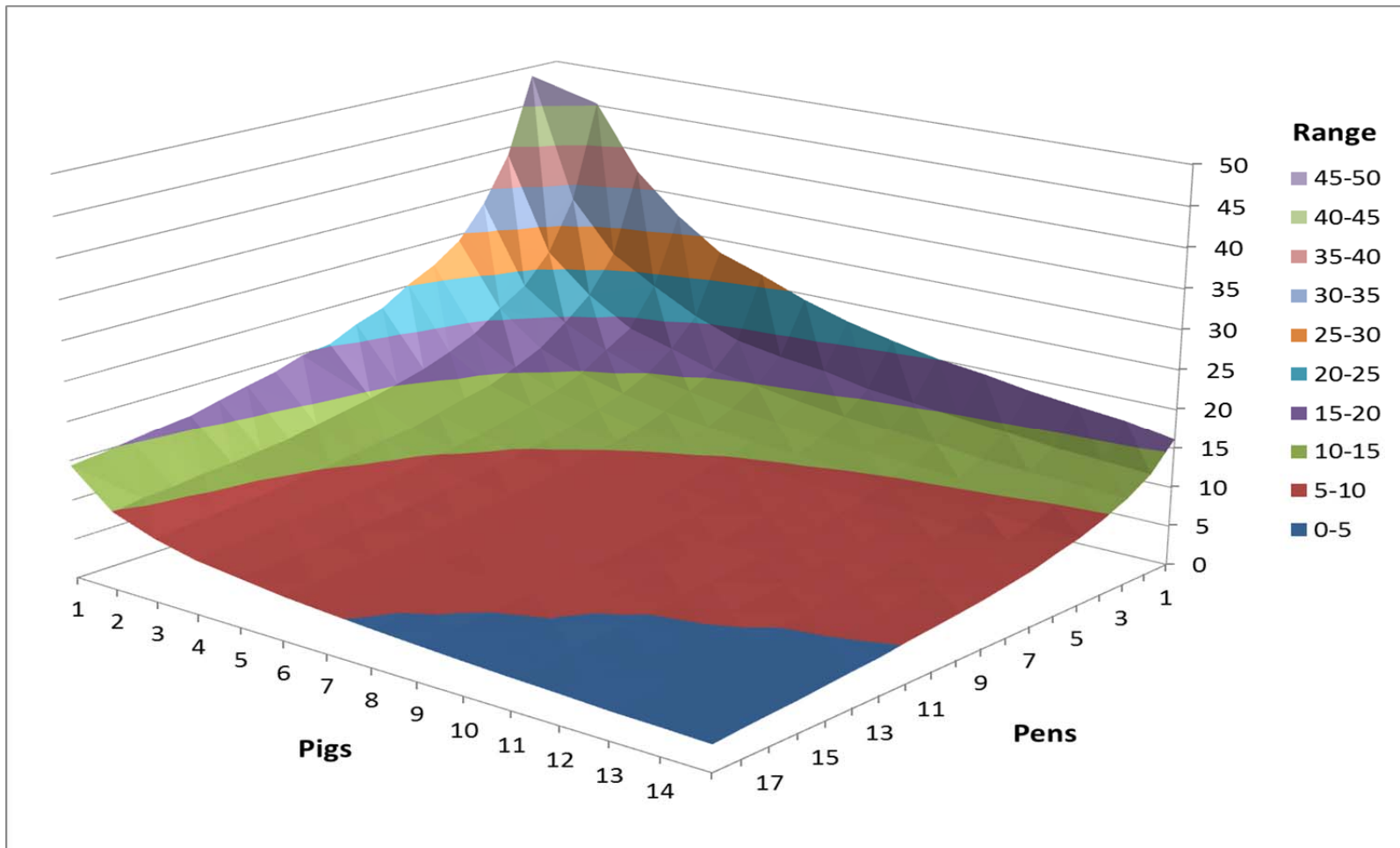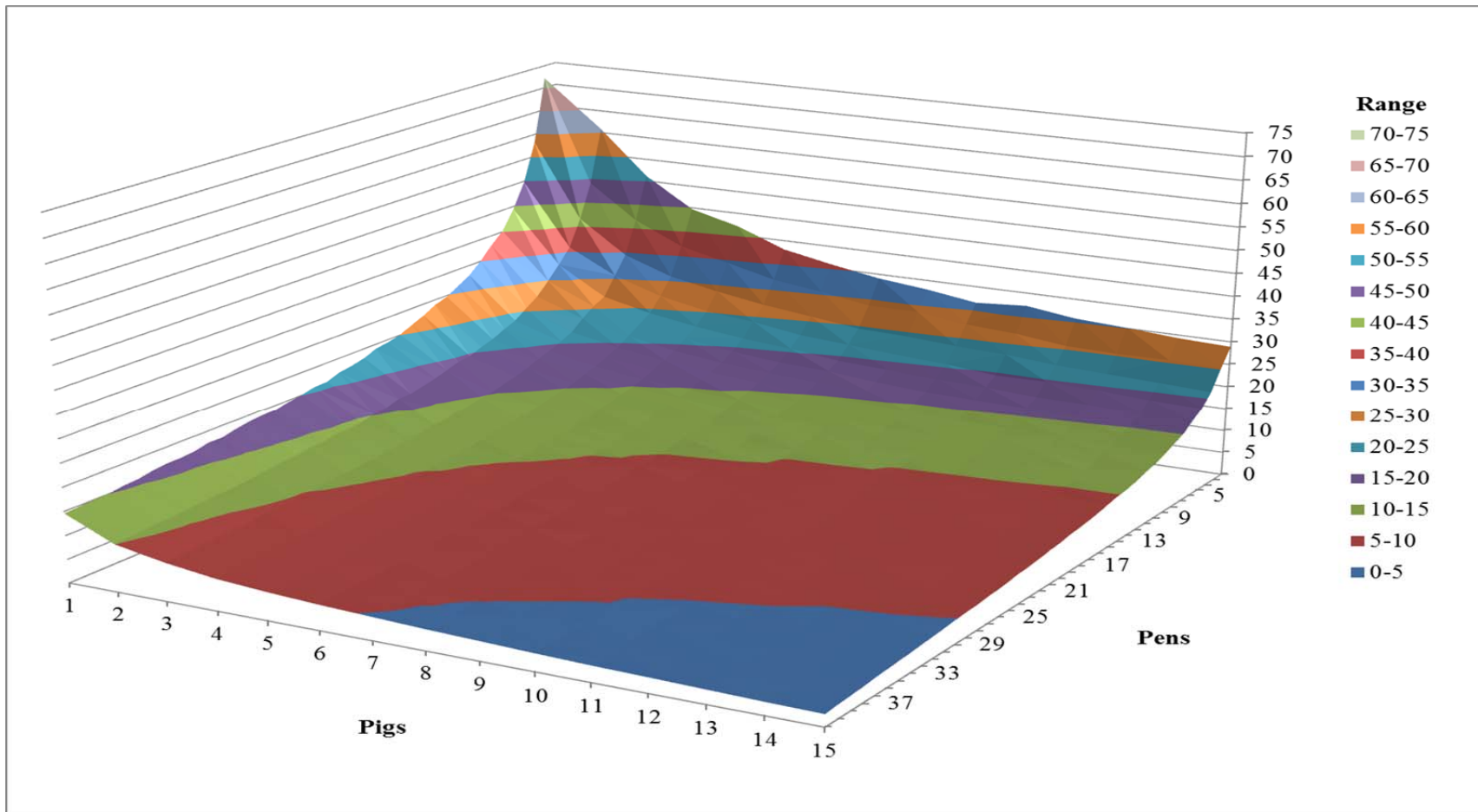
Figure 10. For Dataset C, individual pig weights were collected on a total of 1,069 pigs weighed (population mean = 222.4 lb, median = 224 lb, standard deviation = 32.0 lb, and CV = 14.4%) from 40 pens with 20 to 35 pigs per pen. The dataset was analyzed by altering the number of pigs selected within pens, and total number of pens sampled. This operation was completed 10,000 times for each sampling method, and the range or difference between the upper and lower CI was calculated. Each point on this graph shows the range between the upper and lower CI, represented in pounds.

Table 4. The range between the upper and lower confidence interval (CI) for varying pigs and pen as presented in Figure 10 (dataset C)[1]

Pens, n	Number of pigs from each pen														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		61	51	45	42	38	36	34	33	31	32	31	30	29	29
2	72	49	42	37	34	32	31	29	29	27	27	26	25	26	25
3	58	42	34	31	29	27	26	25	24	24	23	22	22	21	21
4	50	36	31	27	25	24	23	22	21	21	20	20	19	19	18
5	44	33	27	24	23	21	21	19	19	18	18	17	17	17	17
6	40	30	25	23	21	19	18	18	17	16	16	16	16	15	15
7	38	27	24	21	19	18	17	16	16	15	15	15	14	14	14
8	35	26	21	19	18	17	16	15	15	14	14	14	13	13	13
9	33	24	20	18	17	16	15	14	14	14	13	13	13	12	12
10	31	23	19	17	16	15	14	13	13	13	12	12	12	12	12
11	30	21	18	16	15	14	13	13	12	12	12	11	11	11	11
12	29	21	17	15	14	14	13	12	12	11	11	11	11	10	10
13	27	20	17	15	14	13	12	12	11	11	11	10	10	10	10
14	26	19	16	14	13	12	12	11	11	10	10	10	10	9	9
15	25	18	15	14	12	12	11	11	10	10	10	9	9	9	9
16	24	18	15	13	12	11	11	10	10	10	9	9	9	9	9
17	24	17	14	13	11	11	10	10	9	9	9	9	9	8	8
18	23	16	14	12	11	10	10	9	9	9	8	8	8	8	8
19	22	16	13	12	11	10	10	9	9	8	8	8	8	8	7
20	22	15	13	12	10	10	9	9	8	8	8	8	8	7	7
21	21	15	13	11	10	10	9	8	8	8	8	7	7	7	7
22	21	14	12	11	10	9	9	8	8	8	7	7	7	7	7
23	20	14	12	11	9	9	8	8	8	7	7	7	7	7	6
24	20	14	12	10	9	9	8	8	7	7	7	7	6	6	6
25	19	14	11	10	9	8	8	7	7	7	7	6	6	6	6
26	19	13	11	10	9	8	8	7	7	7	6	6	6	6	6
27	19	13	11	9	9	8	7	7	7	6	6	6	6	6	5
28	18	13	11	9	8	8	7	7	6	6	6	6	5	5	5
29	17	13	10	9	8	7	7	6	6	6	6	5	5	5	5
30	17	12	10	9	8	7	7	6	6	6	5	5	5	5	5
31	17	12	10	8	8	7	6	6	6	5	5	5	5	5	4
32	17	12	10	8	7	7	6	6	6	5	5	5	5	4	4
33	17	11	9	8	7	7	6	6	5	5	5	5	4	4	4
34	16	11	9	8	7	6	6	5	5	5	5	4	4	4	4
35	16	11	9	8	7	6	6	5	5	5	4	4	4	4	4
36	15	11	9	8	7	6	5	5	5	4	4	4	4	4	3

37	15	11	9	7	6	6	5	5	5	4	4	4	4	3	3
38	15	11	8	7	6	6	5	5	4	4	4	4	3	3	3
39	15	10	8	7	6	6	5	5	4	4	4	3	3	3	3
40	15	10	8	7	6	5	5	4	4	4	3	3	3	3	3

[1] Colors match the color scheme in Figure 10, representing a range of 5 lb for each color.

Table 5. The resulting mean, upper confidence interval (CI), lower CI, and range when sampling a varying number of pigs and pens to give a total sample size of 30 pigs when estimating the standard deviation of the population

Sampling method	Mean of 10,000 simulations[1]	Upper CI	Lower CI	Range
Dataset A[2]				
15 pigs from 2 pens	32.0	42.9	23.0	19.9
10 pigs from 3 pens	32.2	42.6	23.0	19.6
6 pigs from 5 pens	32.4	42.5	23.2	19.3
5 pigs from 6 pens	32.4	42.3	23.4	18.9
3 pigs from 10 pens	32.5	42.5	23.6	18.8
2 pigs from 15 pens	32.6	42.5	23.5	19.0
1 pig from 30 pens	32.5	42.3	23.8	18.5
Dataset B[3]				
30 pigs from 1 pen	19.8	26.1	14.0	12.1
15 pigs from 2 pens	20.6	27.5	14.6	12.9
10 pigs from 3 pens	20.9	27.9	15.0	12.9
6 pigs from 5 pens	21.1	27.6	15.3	12.3
5 pigs from 6 pens	21.2	27.6	15.3	12.3
3 pigs from 10 pens	21.3	27.5	15.8	11.7
2 pigs from 15 pens	21.4	27.3	15.9	11.4
Dataset C[4]				
15 pigs from 2 pens	29.0	45.0	19.8	25.2
10 pigs from 3 pens	29.9	43.9	20.3	23.6
6 pigs from 5 pens	30.5	42.6	21.2	21.3
5 pigs from 6 pens	30.7	42.3	21.4	20.9
3 pigs from 10 pens	31.1	41.6	22.4	19.2
2 pigs from 15 pens	31.3	41.2	22.8	18.5
1 pig from 30 pens	31.4	40.6	23.1	17.5

[1] The standard deviation was calculated for each of the generated samples, and the mean of the 10,000 generated standard deviation estimates was determined.

[2] A total of 1,260 pigs (mean = 253.0 lb, median = 254 lb, standard deviation = 32.8 lb, and CV = 12.98%) from 48 pens with 23 to 28 pigs per pen.

[3] A total of 1,261 pigs (population mean = 213.5 lb, median = 214 lb, standard deviation = 21.5 lb, and CV = 10.1%) from 19 pens with 56 to 81 pigs per pen.

[4] A total of 1,069 pigs weighed (population mean = 222.4 lb, median = 224 lb, standard deviation = 32.0 lb, and CV = 14.4%) from 40 pens with 20 to 35 pigs per pen.

Table 6. The percentage of the selected pigs as the actual n heaviest or lightest pig[1]

	Rank of pigs					
	1	2	3	4	5	> 5
Heaviest[2]						
Dataset B marketer 1, %	47.4	5.3	0.0	5.3	10.5	31.5
Dataset C marketer 2, %	42.5	35.0	10.0	7.5	5.0	0.0
Dataset C marketer 3, %	55.0	25.0	10.0	2.5	2.5	5.0
Lightest[3]						
Dataset B marketer 1, %	57.9	21.1	10.5	0.0	0.0	10.5
Dataset C marketer 2, %	75.0	17.5	5.0	2.5	0.0	0.0
Dataset C marketer 3, %	68.4	10.5	7.9	5.3	7.9	0.0

[1] Marketers were asked to select the heaviest and lightest pig in each pen in the barn.
[2] 1 is the heaviest pig; 5 is the 5th heaviest pig.
[3] 1 is the lightest pig; 5 is the 5th lightest pig.

Table 7. The resulting mean standard deviation, upper 95% confidence interval (CI), lower 95% CI, and range for the various sampling methods with a total sample size of 30 pigs

Sampling method	Mean standard deviation	Upper CI	Lower CI	Range
Dataset A[1]				
Method 1, 30 random pigs[2]	32.5	42.2	23.5	18.7
Method 2, 2 pigs from 15 pens[3]	32.6	42.5	23.5	19.0
Method 3[4]	32.1	39.2	27.3	11.8
Dataset B[5]				
Method 1, 30 random pigs[2]	21.3	27.3	15.7	11.6
Method 2, 2 pigs from 15 pens[3]	21.4	27.3	15.9	11.4
Method 3[4]	22.8	24.2	19.8	4.3
Dataset C[6]				
Method 1, 30 random pigs[2]	31.7	41.4	23.2	18.2
Method 2, 2 pigs from 15 pens[3]	31.3	41.2	22.8	18.5
Method 3[4]				
Marketer 2	32.2	40.3	23.5	16.8
Marketer 3	32.3	40.3	23.8	16.5

[1] A total of 1,260 pigs (mean = 253.0 lb, median = 254 lb, standard deviation = 32.8 lb, and CV = 12.98%) from 48 pens with 23 to 28 pigs per pen.

[2] 30 pigs were randomly selected from the barn.

[3] 2 random pigs were selected from 15 randomly selected pens.

[4] Select the heaviest and lightest pig (determined visually) in each pen, subtract the lightest weight from the heaviest weight, and divide by 6.

[5] A total of 1,261 pigs (population mean = 213.5 lb, median = 214 lb, standard deviation = 21.5 lb, and CV = 10.1%) from 19 pens with 56 to 81 pigs per pen.

[6] A total of 1,069 pigs were weighed (population mean = 222.4 lb, median = 224 lb, standard deviation = 32.0 lb, and CV = 14.4%) from 40 pens with 20 to 35 pigs per pen.