

Yield Probability Distribution Analysis: A Forgotten Tool In The Farm Decision Toolkit?

Doug Wilcox, Manitoba Agricultural Services Corporation – Insurance Unit, 400-50-24th Street N.W.,
Portage La Prairie, Manitoba, R1N 3V9. Email: dwilcox@masc.mb.ca

Abstract: This paper examines the issue of utilizing crop-yield distribution information for better on-farm decision making. Yield distribution variables are calculated using Manitoba data and the results demonstrate that Manitoba yield distributions are often negatively skewed. The advantage of recognizing non-normal distributions and not assuming normality is explored. It is also demonstrated how yield distribution variables can be easily calculated using spreadsheet formulas and readily available data..

Introduction

Farmers are faced daily with the difficult job of making decisions that require them to come up with realistic estimates of future yields. For example realistic yield estimates are critical when pencilling out which varieties to select or how much fertilizer to purchase. A yield that is too aggressive can result in increased cost and a yield that is too conservative will minimize profit potential.

My experience is that many farm decision makers have a bias – a tendency to ignore low crop yields and focus on higher crop yields. This bias has been observed by others as well (Clon-Gallart and Juarez-Rubio, 2005). Farm decision makers who are savvy enough to realize that a biased “guesstimate” approach isn't good enough for important decisions are likely using "average" yields from their farm or region for their realistic estimate. Although better than a guesstimate approach, there are still limitations to using "average" yields. To overcome these limitations and to assist with making better decisions, farm decision makers should be taking into account crop-yield probability distributions.

Crop-yield probability distributions have been extensively studied, mainly in regards to premium ratemaking for crop insurance programs (Day, 1965; Yeh and Sun, 1980; Ramerez et al, 2003) and more recently for use in farm decision support systems (Clon-Gallart and Juarez-Rubio, 2005). The results of these studies can be exploited by farm decision makers to enhance the outcomes of their own yield related decisions.

Some Background Statistics

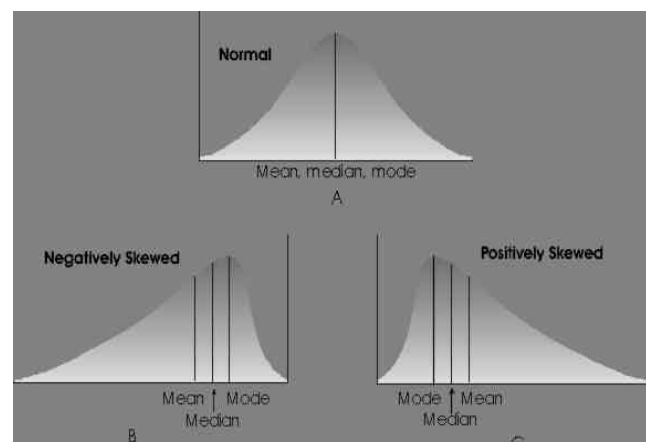
Before exploring why knowledge of crop-yield probability distributions will help with making better farm decisions it would likely be useful to give a quick review some of the simple statistics of yield distributions that will referred to in this paper.

Probability Distribution - probability and relative frequency can be used interchangeably, so a yield probability distribution is simply a plot of how often you can expect a particular yield to occur.

Mean (Average) – The arithmetic mean of a variable – the sum of all the values divided by the number of (nonmissing) values.

Median – The median estimates the center of the distribution. If the data are sorted in increasing order the median is the value above which half the values fall.

Mode – The value that appears most frequently



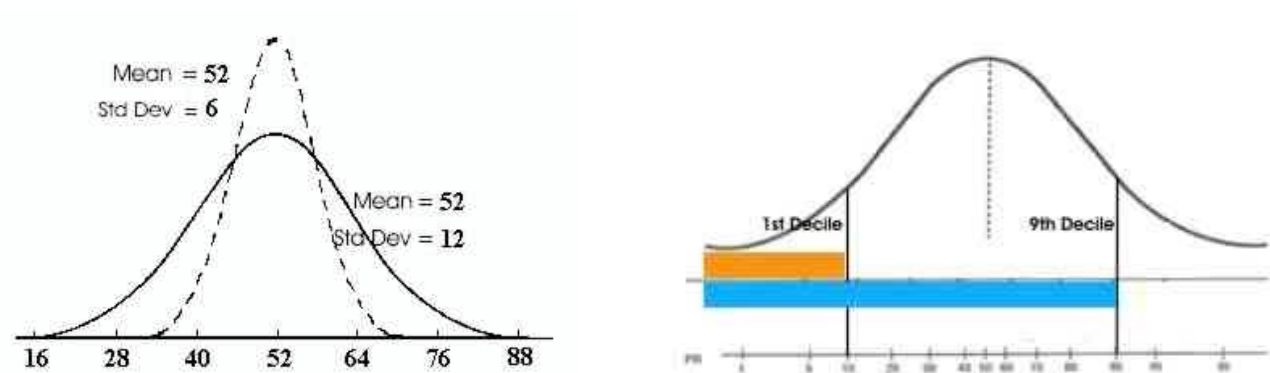
Skewness – A measure of the symmetry of a distribution about its mean. If skewness is significantly nonzero, the distribution is asymmetric. A significant positive value indicates a long right tail (positively skewed); a negative value, a long left tail (negatively skewed).

Normal Distribution – A bell shaped frequency distribution where the mean, median and mode all have the same value.

Standard Deviation – A measure of spread of the data. It is the square root of the variance

Coefficient of Variation (C.V.) – A scale adjusted measure of the spread of data. It is the ratio of the standard deviation divided by the mean.

Percentiles – A measure of what percentage of data is at or below a given value. The 50th percentile is called the median. A true percentile splits the distribution of an ordered set of data into hundredths. Deciles split the ordered data into tenths and quartiles split the ordered data into quarters.



These basic distribution statistics can all be calculated easily using a computer and a spreadsheet program such as Microsoft Excel. The relevant commands to apply in Excel are listed in Appendix 1 to this paper for your information.

What's Wrong With Crop-Yield Averages?

Quite often yields are non-normally distributed with a tendency towards skewness (Day, 1965; Yeh and Sun, 1980; Ramirez et al, 2003). The issue of normality of data versus non-normality of data is important. For example when you determine an average (mean) yield, whether you realize it or not, you are making a distribution assumption of normality. You may think that the crop-yield average you have calculated is the mid-point yield but since most yield distributions are non-normal it is not likely to be the mid-point yield. This incorrect mid-point yield assumption could fundamentally impact your farm decisions in a negative fashion. There are statistical techniques to transform non-normal data to normal – but these are for more advanced users.

Historically crop average yields have been assumed to represent the mid-point value of a crop-yield probability distribution because it has been the simplest to calculate. Additionally, given all the other potential errors in the data, in many instances using an average for the mid-point estimate may not be that far off. However, I feel that this is a mistake. Why add to your decision risk by adding in the often erroneous assumption of normality for the sake on computational convenience – especially when generally readily available spreadsheet tools mean that computational convenience is no longer an issue.

Normality is rare with crop-yield data because crop-yields are not truly “independent” because of the multi-field impact of management and environmental influences. Lack of independence between individual crop-yield losses means that the Central Limit Theorem of statistics cannot be used as the basis for assuming crop-yield normality.

If crop-yield averages are suspect, what should you use instead? My experience with Manitoba crop-yield distributions has shown me that if farm decision makers want to maximize profit and minimize risk they must use more than just average crop-yields in their decision making process. To make rational decisions farm decision makers should account for all the likely situations they could face, starting with the true mid-point yield but also with consideration of the variance and skewness of crop-yields.

The best way to illustrate this is to look at some real examples. Data for all examples in this presentation come from data in the Manitoba Agricultural Services Corporation (MASC) database. The yields are the individual farm variety by crop yields for the period 1996 to 2004, inclusive. The yields analyzed are the net yields (i.e. yields have been adjusted for dockage and moisture, but not quality).

Comparing Crop-Yield Variability (Risk)

Farm decision makers should not make decisions or plans based exclusively on the assumption of average conditions. To make rational decisions they must account for the risk situations they are likely to face, starting with consideration of yield variation. For example, crop yield variation is the cause of crop insurance claims. If a farm decision maker’s crop consistently has little or no yield variation then they will have few losses - the reverse is also true. A farm decision maker should therefore increase crop insurance coverage as yield variability increases.

Traditionally, variance and standard deviation has been used as measures of data variability. A problem with variance and standard deviation is that they are difficult to interpret without knowing the level or magnitude of the underlying variable. A variance of 10 bushels has quite different implications for the distribution when the yield is 30 bushels per acre than when it is 100 bushels per acre. Therefore, an alternative measure of variability that takes into account proportionality called the Coefficient of Variation (C.V.) is commonly used to facilitate comparison .

For example if we compare the C.V. of the crops listed in Table 1 we find the following over the period 1996 to 2004 in Manitoba:

Canola<Red Spring Wheat<Winter Wheat<Barley<Flax<Field Peas<Sunflowers<Navy Beans<Oats<Grain Corn<Buckwheat<Soybeans.

Canola has a C.V. of 34% and Soybeans has a C.V. of 70%. If we assume that yield variability equates to “riskiness” then in Manitoba argentine canola and red spring wheat would be the lowest risk crops and buckwheat and soybeans would be the highest risk crops. The only surprise to me in this listing is the placement of oats with the higher risk crops – this may be an artifact of a large yield range and a higher than average zero yield listing (for a cereal).

Figures A to F illustrate the yield distributions for some selected crops. One of the most obvious differences between the higher risk crops and the lower risk crops is that the higher risk crops have more zero yields. The other characteristic to note is that all the distributions are negatively skewed.

Table 2 presents the yield distribution statistics for argentine canola grown in the various risk areas of Manitoba . the C.V. values range from 24% for risk area 16 to 56% for risk area 14. There are no clear regional trends other than the general one that the regions with “heavier” soils tend to be the highest risk

regions for canola. Additionally, it generally appears that the western half of the province has less yield variability than the eastern half of the province.

The C.V. values for the most popular two varieties of some selected major Manitoba are listed in Table 3. The results indicate that varieties can differ significantly in C.V. values. Differences in variety yield stability as indicated by C.V. values might be worth considering when making variety choices. Certainly when all other data for varieties you wish to compare are similar, the variety with lower variability would be the less risky choice. Additionally, maybe purely coincidentally, in this data the highest yielding varieties also had the lowest C.V. values. Figures G to J illustrate the differences in variety distributions for two white pea bean varieties and two flax varieties.

Table 4 list the C.V. values for several crops for the two contrasting years 2003 and 2004. The 2003 crop year was noted for record breaking high yields whereas 2004 was one of the coldest growing seasons on record, generally resulting in poor yields for long season crops. This is reflected in the fact that the C.V. values increased significantly in 2004, compared to 2003, for the long season crops but much less so for the shorter season crops. The differences between 2003 and 2004 also show up in frequency charts illustrated in Figures K to R.

As a farm decision maker, having crop-yield variability information can provide additional information that will help with making a more informed decision. In particular, the utility of varietal C.V. differences has been underutilized by farm decision makers..

Median Versus Mean (Average)

When planning or decision making you must determine your objective. If your objective is to maximize the “odds” of coming close to the observed values then you should target the mode. If the objective is to have a nearly equal chance of over and under-predicting then you should target a median. The average or mean is often used for the same objective as the median, but truly only applies as such when the data is normally distributed. Median estimates of yield are preferred to mean estimates because they reflect the 50% percentile, whereas the mean may not.

Table 1 lists a comparison of the mean, median and mode values for various Manitoba crops. It is interesting that the mode for most of the higher risk crops is zero. This is mainly the result of the impact of 2004. Also it is noteworthy that except for buckwheat the median is larger than the mean, implying that the crop-yield distributions are negatively skewed. Figures C and D illustrate why the modes for grain corn and soybeans are zero. Similarly Tables 2 to 4 also have median>mean statistics, indicating that crop-yields are negatively skewed. Yeh and Sun, 1980, similarly found that wheat yield distributions in Manitoba were often negatively skewed. Although not explored in this paper, other researchers such as Day (1965) have found that as more technological improvements are made (e.g. more fertilizer, irrigation) that yields become more positively skewed (less negatively skewed).

When setting a yield target the goal is generally to have a nearly equal chance of over and under-estimation. This means for non-normal distributions that a farm decision maker should target the median – not the average.(mean). Although in many cases the difference between the mean and median values was small (e.g. 1 bushel) this could result in a difference when pencilling out scenarios that means the acceptance or rejection of a particular course of action – especially when combined with price and acreage information.

Best Case/ Worst Case – Use of Percentiles

When farm decision makers are planning they should be aware of not only the average situation but also be prepared for alternative situations. Producers should use explore the consequences of a range of yields and plan for best and worse case situations. Farm decision makers should develop a management strategy aimed towards achieving the best case but which ensures preparedness for a worse case. For this kind of analysis I would propose that farm decision makers use the 1st decile as their worst case (the yield at which 10% of the time yields are lower) and the 9th decile as their best case (the yield at which 90% of the time yields are lower).

When comparing the 1st and 9th deciles for the major crops in Table 1 it stands out that the 1st decile values are at or near zero for the crops soybean, grain corn, and buckwheat. In other words when these crops are grown in Manitoba one can expect to have zero yields at least one out of ten years – maybe that is why we have nearly 100% crop insurance sign-up at maximum coverage for these crops! In contrast, in the case of the 9th decile, a farm decision maker may want to make sure they have enough storage capacity for a 114 bushel oat crop, or enough drying capacity for a 126 bu/ac grain corn crop. It is also interesting to note that the ninth decile values are similar to the statistical extreme values for Manitoba listed by Wilcox (2004b).

If a farm decision maker doesn't have the necessary data, or doesn't want to do the necessary statistical calculations required to determine the 1st decile and the 9th decile, applying the percentages in the following table to a personal median crop-yield should give a close approximation. For example if you are trying to estimate the 1st and 9th decile for a flax crop you would take 41% and 145% of the median, respectively.

Crop	1st Decile	Median	9th Decile	Units
	As % Of Median		As % Of Median	
Red Spring Wheat	56%	37.7	140%	bu/ac
Winter Wheat	52%	56.4	138%	bu/ac
Barley	43%	59.0	141%	bu/ac
Oats	29%	71.5	159%	bu/ac
Argentine Canola	54%	30.3	138%	bu/ac
Flax	41%	19.6	145%	bu/ac
Field Peas	40%	35.2	147%	bu/ac
Soybeans	0%	21.1	177%	bu/ac
Grain Corn	0%	91.4	136%	bu/ac
Buckwheat	14%	13.8	208%	bu/ac
White Pea Beans	25%	1461.6	149%	lbs/ac
NonOil Sunflowers	30%	1380.1	146%	lbs/ac

Applying in The Real World – Close Approximations

I appreciate that using crop-yield probability distributions based on historical observations is not necessarily the best estimator of future yield. Crop-yield probability distributions based on historic observations assume that average weather conditions will prevail between planting and harvest. If the growing season is poorer than usual or better than usual than the yield will be expected to correspondingly change. The best estimates of future yield will be obtained if the farm decision-maker takes into consideration location, individual management, and field history. Trends also come into play - the ballpark trend for Manitoba crop yields over the period 1974 to 2002 is a doubling of yields (Wilcox, 2004a). However, in the absence of these field specific records which are adjusted for trends, regional and even provincial records are superior to a "guesstimate". Particularly when such regional yields are

adapted to local conditions on the basis of logic and experience. A potential source of good Manitoba crop yield data by region is of the Manitoba Management Plus web site (<http://www.mmpp.com/>). Yield data by crop variety, fertilizer and pesticide use can be found there.

Keeping records on all fields should be normal practice in any farming operation. These records will pinpoint problems, enhance decision making, and possibly provide information on better management. MASC annually provides a Crop Management History report to its clients. This report contains the last five years of cropping history by field and crop. This data can be used to establish farm specific crop yield distributions. If five years of data is not enough MASC clients can visit their local MASC agent to have a report with more years of data generated.

A farm decision maker can then input their own field history records or the Crop Management History report into a spreadsheet and generate crop-yield distributions specific to their farm. Appendix 1 has a description of the necessary spreadsheet calculations and an example spreadsheet outline. The farm decision maker that creates or obtains this kind of spreadsheet can have farm specific statistics - median, 1st decile, 9th decile and measures of variability at their fingertips– all of which can be used to make better farm management decisions.

Summary Remarks

The main goal of this presentation is not to turn farm decision makers into statisticians but is to remind them of the value of the analysis of crop-yield distributions for decision making. Additionally, farm decision makers need to recognize that crop-yield distributions may be non-normal which will require non-normal methodology to generate useful statistics for decision making. These non-normal statistical tools are not new but due to data unavailability and data analysis difficulties these tools had been forgotten at the bottom of the farm decision maker's toolbox. No more - Manitoba and farm specific data is now readily available and with the power of readily available spreadsheets these useful statistics are easily calculated, even by occasional number crunchers.

References:

Clop-Gallart, M. M. and Juarez-Rubio, F., 2005. Elicitation of subjective crop yield PDF for DSS implementation. Paper presented at the 11th seminar of the EAAE, August 24-27, 2005. 16 pgs.

Day, R.H. 1965. Probability distributions of field crop yields. *J. Farm Economics*. 47: 733-740.

Ramirez, O.A., Misra, S. and Field, J. 2003. Crop-yield distributions revisited. *Amer. J. Agr. Econ* 85:108-120.

Wilcox, D. 2004a. The forgotten success – Manitoba yield increases. pgs. 10-11, in *Yield Manitoba 2004*, Supplement to the *Farmers Independent Weekly* February 19, 2004.

Wilcox, D. 2004b. Extreme yields in Manitoba. If everything goes right, how high can yields go? pg. 12 in *Yield Manitoba 2004*, Supplement to the *Farmers Independent Weekly* February 19, 2004.

Yeh, M. H. and Sun, S. 1980. An evaluation of the actuarial structure of Manitoba Crop Insurance program and its implications. *Can. J. Ag. Econ*. 28:54-70.

Table 1: Yield distribution statistics for various Manitoba crops. Based on farm yields reported to MASC over the period 1996 to 2004.

Crop	Units	N of Cases	Mean Yield	Median Yield	Mode	Standard Deviation	C.V.	Skewness	1st Decile	9th Decile
Red Spring Wheat	bu/ac	89,502	37.2	37.7	40.4	12.8	0.344	-0.117 *	21.0	52.9
Winter Wheat	bu/ac	6,172	54.6	56.4	60.6	19.0	0.347	-0.503 *	29.3	77.6
Barley	bu/ac	52,598	56.4	59.0	59.7	22.6	0.401	-0.402 *	25.3	83.4
Oats	bu/ac	44,544	69.8	71.5	0.0	34.9	0.500	-0.110 *	21.0	113.7
Argentine Canola	bu/ac	104,770	29.6	30.3	29.5	10.1	0.343	-0.429 *	16.3	41.8
Flax	bu/ac	28,702	18.9	19.6	19.7	7.8	0.416	-0.303 *	8.0	28.4
Field Peas	bu/ac	8,707	34.2	35.2	40.4	14.4	0.421	-0.267 *	14.1	51.7
Soybeans	bu/ac	3,951	19.8	21.1	0.0	13.7	0.696	0.049	0.0	37.4
Grain Corn	bu/ac	8,242	78.6	91.4	0.0	43.7	0.555	-0.660 *	0.0	124.6
Buckwheat	bu/ac	2,055	14.7	13.8	0.0	10.0	0.683	0.556 *	1.9	28.7
White Pea Beans	lbs/ac	5,560	1375.7	1461.6	0.0	668.0	0.486	-0.333 *	363.8	2180.3
Non-Oil Sunflowers	lbs/ac	4,852	1302.9	1380.1	1543.2	601.9	0.462	-0.335 *	416.7	2017.2

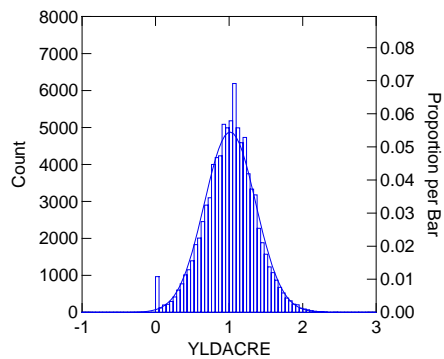


Figure A: Wheat Yield Distribution

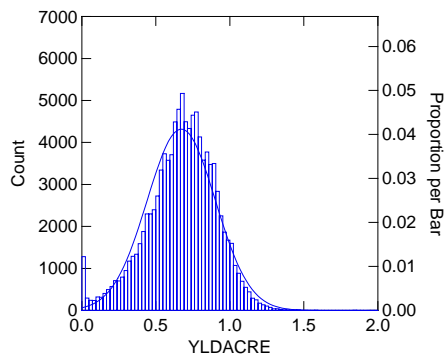


Figure B: Canola Yield Distribution

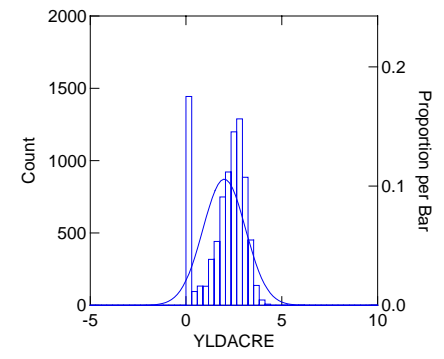


Figure C: Grain Corn Yield Distribution

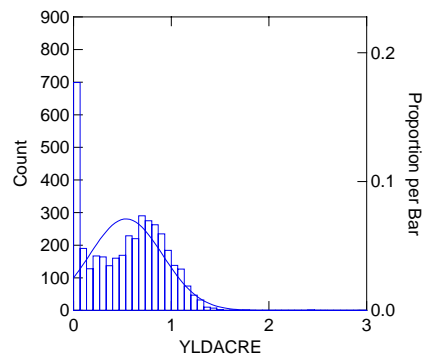


Figure D: Soybean Yield Distribution

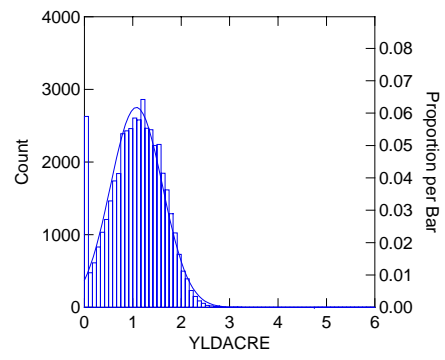


Figure E: Oat Yield Distribution

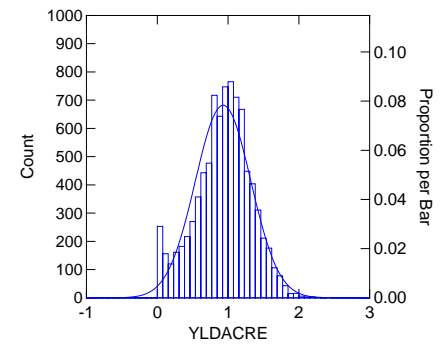


Figure F: Field Pea Yield Distribution

Figures A to F: Yield (tonnes/ac) distribution frequency charts for selected crops grown in Manitoba. Based on farm yields reported to MASC over the period 1996 to 2004. For comparison purposes the expected normal distribution is shown as a solid spline.

Table 2: Yield (bu/ac) distribution statistics for Argentine Canola grown in various risk areas of Manitoba. Based on farm yields reported to MASC over the period 1996 to 2004.

Risk Area	N of Cases	Mean Yield	Median Yield	Standard Deviation	C.V.	Skewness
1	2,804	22.1	22.5	8.1	0.368	-0.296 *
2	8,924	28.0	28.8	8.1	0.290	-0.491 *
3	3,748	24.7	25.2	8.7	0.352	-0.374 *
4	6,424	27.9	28.3	8.9	0.318	-0.318 *
5	12,792	34.2	34.7	8.8	0.259	-0.571 *
6	10,107	27.6	28.5	8.9	0.321	-0.663 *
7	4,815	27.2	28.5	9.4	0.346	-0.646 *
8	5,377	30.5	31.0	9.6	0.314	-0.415 *
9	10,033	28.5	29.4	9.7	0.342	-0.500 *
10	2,542	29.4	30.0	10.6	0.362	-0.407 *
11	7,102	31.8	32.8	10.1	0.318	-0.603 *
12 - 12 Soils	4,835	34.0	34.9	9.7	0.285	-0.538 *
12 - 32 Soils	17,491	31.0	32.5	11.0	0.355	-0.475 *
14 - 14 Soils	211	21.8	22.6	12.3	0.563	-0.053
14 - 34 Soils	3,201	28.2	29.5	12.1	0.427	-0.334 *
15 - 15 Soils	103	27.0	28.2	10.8	0.401	-0.468
15 - 35 Soils	2,857	29.9	30.7	11.2	0.373	-0.340 *
16	426	31.3	32.0	7.4	0.237	-0.514 *

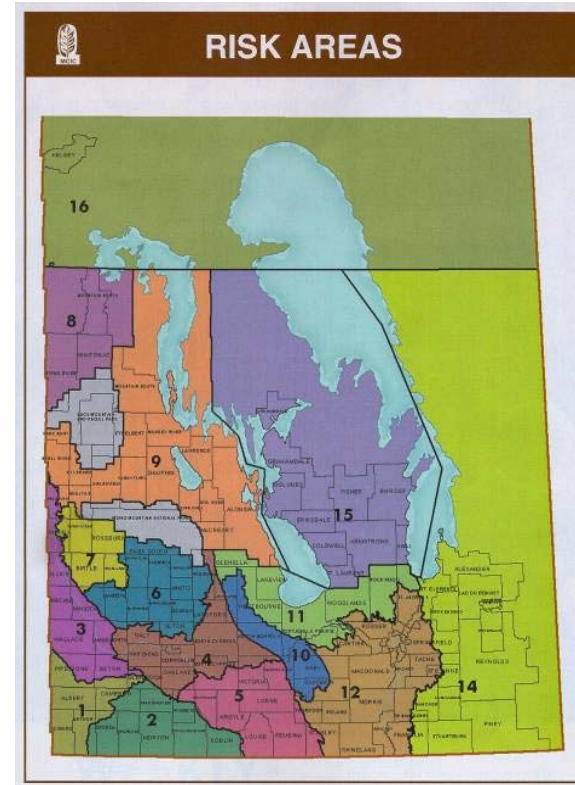
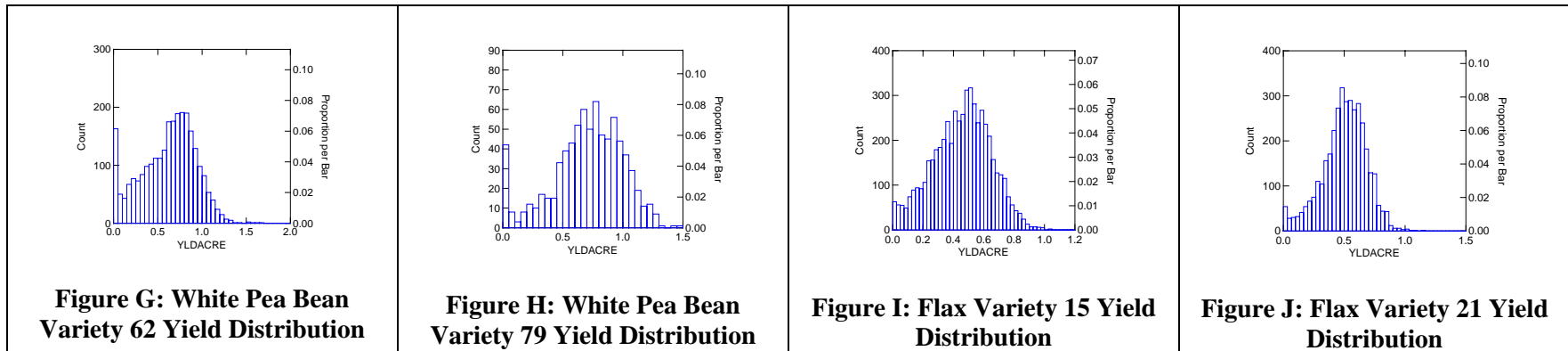


Table 3. Yield distribution statistics for the two major varieties of selected crops grown in Manitoba. Based on farm yields reported to MASC over the period 1996 to 2004.

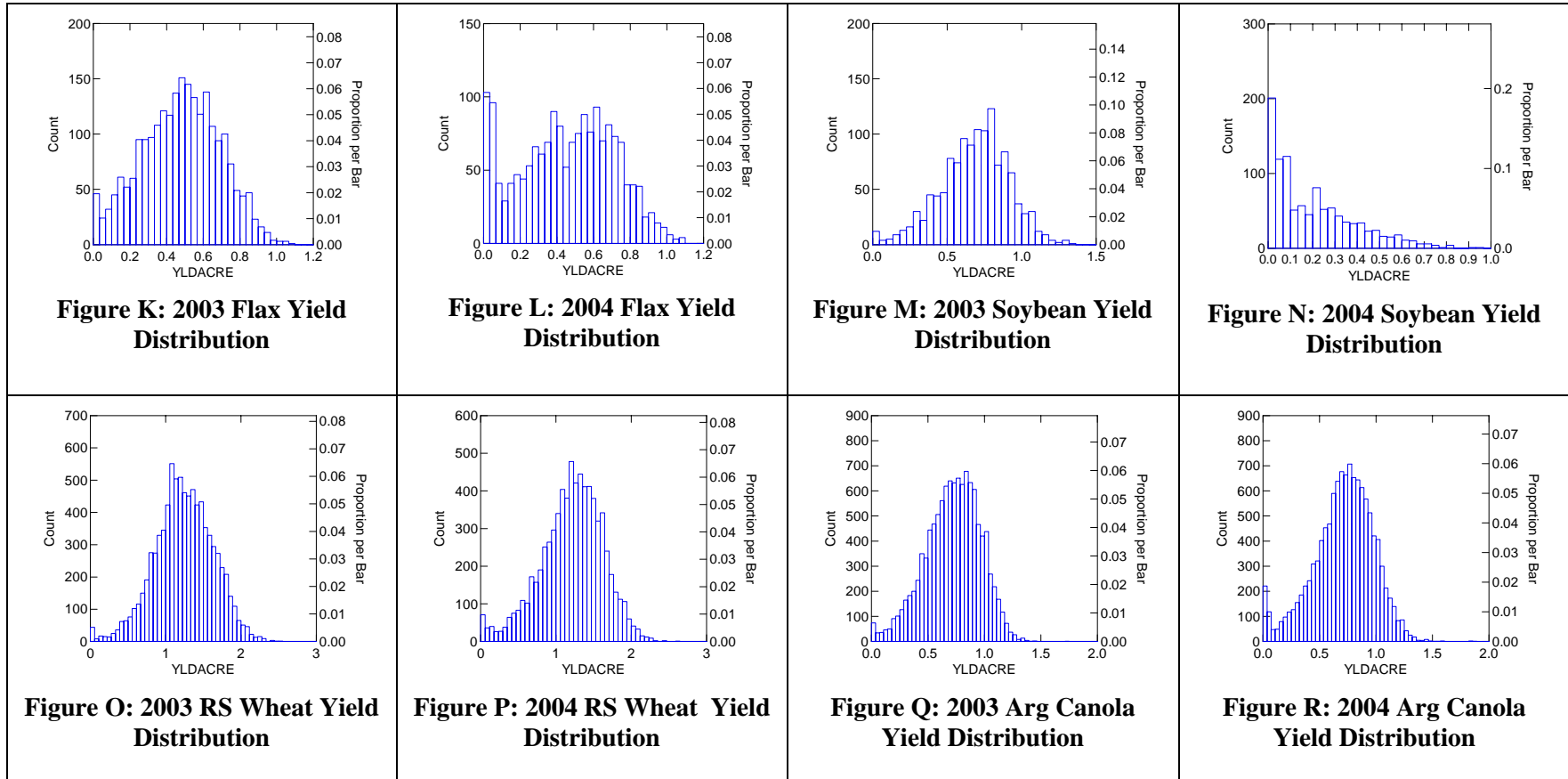
Crop	Variety	Units	N of Cases	Mean Yield	Median Yield	Standard Deviation	C.V.	Skewness
Red Spring Wheat	20	bu/ac	19487	37.3	38.0	11.5	0.307	-0.158*
Red Spring Wheat	31	bu/ac	33715	38.5	39.0	13.2	0.342	-0.108*
Argentine Canola	85	bu/ac	5495	27.8	28.6	9.1	0.328	-0.429*
Argentine Canola	33	bu/ac	6462	32.1	33.2	9.7	0.304	-0.648*
Flax	15	bu/ac	5411	17.9	18.5	7.5	0.419	-0.170*
Flax	21	bu/ac	3720	19.8	20.2	7.0	0.356	-0.425*
Soybeans	69	bu/ac	600	23.3	25.0	13.5	0.578	-0.186
Soybeans	53	bu/ac	667	18.2	19.4	12.2	0.670	0.089
Grain Corn	232	bu/ac	1543	76.2	98.0	51.2	0.671	-0.530*
Grain Corn	231	bu/ac	751	79.6	95.5	47.4	0.596	-0.708*
White Pea Beans	62	lbs/ac	2647	1360.2	1455.0	676.8	0.498	-0.286*
White Pea Beans	79	lbs/ac	781	1552.0	1616.0	657.0	0.424	-0.560*



Figures G to J: Some yield (tonnes/ac) distribution frequency charts for the two major bean and flax varieties grown in Manitoba. Based on farm yields reported to MASC over the period 1996 to 2004

Table 4. Yield distribution statistics for selected crops grown in Manitoba during the contrasting years 2003 and 2004. Based on farm yields reported to MASC.

Crop	Year	Unit	N of Cases	Mean Yield	Median Yield	Standard Deviation	C.V.	Skewness
Red Spring Wheat	2003	bu/ac	8,530	45.6	45.6	14.2	0.311	-0.199 *
Red Spring Wheat	2004	bu/ac	7,277	44.5	45.6	15.1	0.341	-0.434 *
Argentine Canola	2003	bu/ac	11,358	31.7	32.6	10.7	0.337	-0.367 *
Argentine Canola	2004	bu/ac	11,811	31.4	32.5	11.8	0.374	-0.444 *
Flax	2003	bu/ac	2,350	18.8	19.1	8.4	0.446	-0.077
Flax	2004	bu/ac	1,762	17.9	18.5	10.2	0.571	-0.06
Soybeans	2003	bu/ac	1,263	25.2	25.8	8.5	0.336	-0.342 *
Soybeans	2004	bu/ac	1,068	7.5	5.6	6.9	0.915	1.009 *
Grain Corn	2003	bu/ac	1,395	98.5	106.2	32.8	0.333	-1.117 *
Grain Corn	2004	bu/ac	1,150	1.5	0.0	6.1	3.918	6.997 *
White Pea Beans	2003	lbs/ac	605	1604.9	1660.0	502.6	0.313	-0.341 *
White Pea Beans	2004	lbs/ac	588	361.6	189.6	418.9	1.157	1.032 *



Figures K to R: Yield (tonnes/ac) distribution frequency charts for a few crops grown in Manitoba during the contrasting years 2003 and 2004. Based on farm yields reported to MASC.

Appendix 1

Calculating Crop Yield Distribution Statistics In MS Excel:

Professional statisticians and data analysts usually have powerful software at their disposal. However, if you are reading this the odds are that you are not a professional statistician and that all you need is a simple way to conduct yield probability distribution analysis on software that you already have. You likely already have that capability - most spreadsheet packages for PC have the capability of making yield probability distribution analysis easy. This examples presented here are for Microsoft Excel but will likely work with little or no modifications in other spreadsheet packages as well.

Microsoft Excel's average function is =AVERAGE(range of data)

Microsoft Excel's median function is =MEDIAN(range of data)

Microsoft Excel's mode function is =MODE(range of data)

Microsoft Excel's standard deviation function is =STDEV(range of data)

Microsoft Excel's percentile function is =PERCENTILE(range of data,0.10), to obtain the 10th percentile (1st Decile); to obtain the 9th decile use 0.90 in place of 0.10.

There is no automatic coefficient of variation (C.V.) function in MS Excel but it can be calculated as follows: $C.V. = \text{STDEV}(\text{range of data}) / \text{AVERAGE}(\text{range of data})$

Plotting Crop Yield Distribution Histograms In MS Excel:

You can plot simple plots of crop yield distributions using the Histogram function of MS Excel's Data Analysis ToolPak. Not all Excel users are familiar with Data Analysis ToolPak because it is an Add-In which doesn't appear by default. To load the data Analysis ToolPak, select Tools and then Add-Ins. Click the appropriate box and select OK. After a few seconds, the Add-In will be loaded. To activate the Data Analysis ToolPak, select Tools and then Data Analysis. A pop-up window will appear listing all of the data analysis functions.

Once the Add-In is installed make a list of "bins" or yield categories that you want to be plotted (e.g. 0, 10, 20, 30, etc). Then go to the menu bar and select Tools then scroll down and select Data Analysis. Scroll down, highlight Histogram, then click OK. A pop-up window will appear. In the field titled Input Range, enter the cell range containing the data you wish to plot (e.g. select A2 through A21). In the Bin Range field, enter the cell range containing the list of all category codes, or "bins," you want to be tabulated and charted (e.g. B2 through B10). Next, select where you would like Excel to deposit the finished output. Finally, check the Cumulative Percentage box at the bottom of the pop-up window When you click OK, Excel will generate a tabular summary and a histogram chart of your data. The histogram chart can then be stretched using standard desk top publishing techniques to increase its size and visibility.

Excel's statistical capabilities are quite impressive a easily useable for crop yield distribution analysis. With a little effort, even the occasional number cruncher can fulfill their crop yield distribution analysis needs.

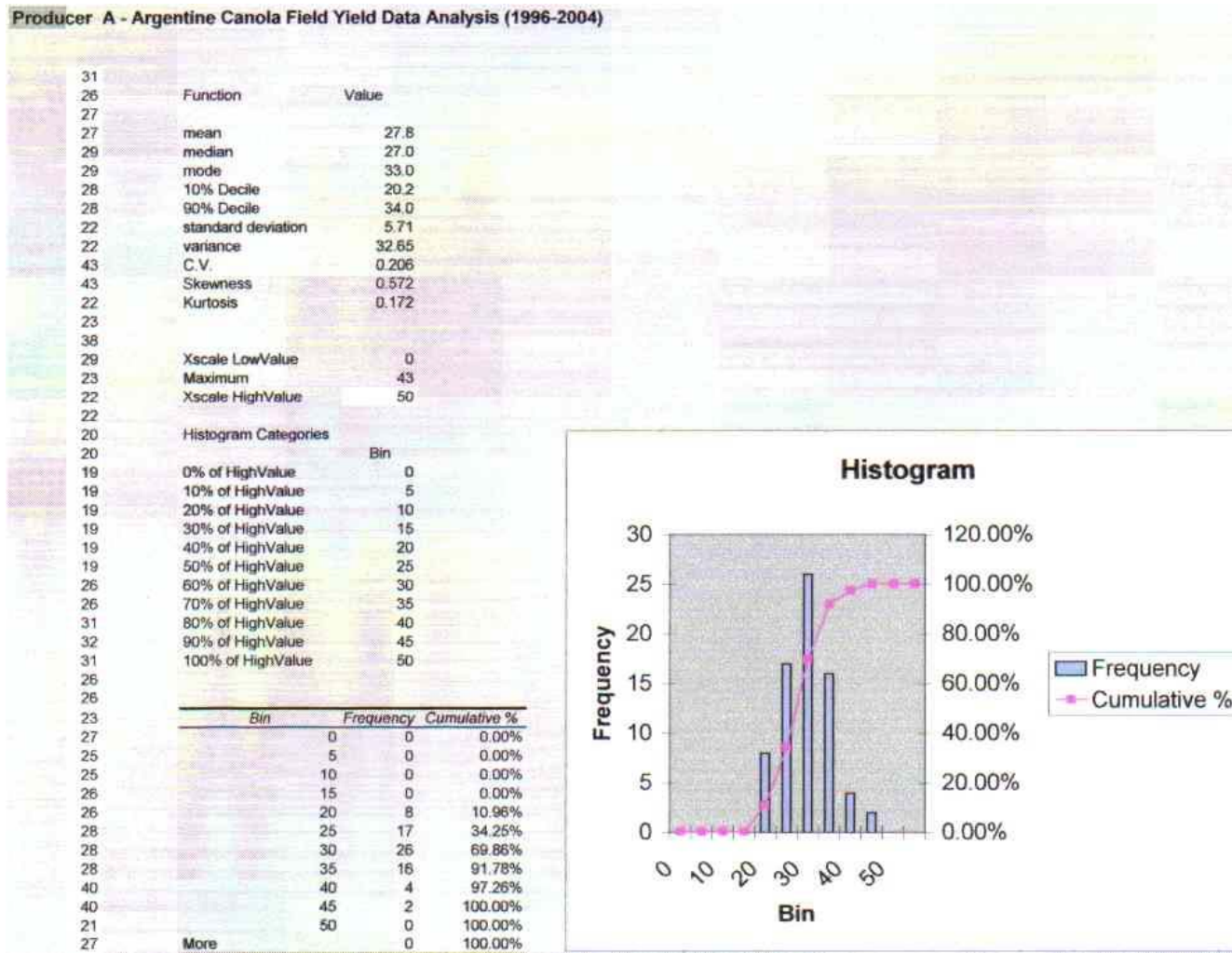


Figure S: Appendix 1 - A sample spreadsheet illustrating the use of MS Excel for crop yield distribution analysis on data from an individual producer.