भारतीय प्रबंध संस्थान बेंगलूर
INDIAN INSTITUTE OF MANAGEMENT BANGALORE

WORKING PAPER NO: 423

Impact of Fuzziness in Measurement Scale on basic Statistical Inference

Shubhabrata Das
Professor
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 5600 76
Ph: 080-26993150
shubho@iimb.ernet.in

Year of Publication-September 2013

IMPACT OF FUZZINESS IN MEASUREMENT SCALE ON BASIC STATISTICAL INFERENCE

Shubhabrata Das

Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore 560076 India

Abstract

When a product, performance or service is rated or a behavioural response is sought, the alternatives typically do not have well-defined and universally understood demarcations. In this research, we address validity or inaccuracies in basic statistical inference based on such fuzzy data. In particular, we focus on inference on population mean and variance of single population as well as two population one-sample and also address the proportion problem in the paired test framework. In the testing of hypothesis framework, both the size and power of the tests are looked at. The results are mixed, as we observe that fuzziness of data impacts the inference substantially in some problems, while it has virtually no impact in some other problem domain.

Keywords: fuzzy, mean, one/two sample problem, power, proportion, scale, size, testing of hypothesis, variance.

1 Introduction

Many survey data is collected in Likert scale and hence most statistical analysis that routinely treats the data as though in interval or ratio scale could be erroneous. Part of the problem often lies with fuzziness of the data due to scale. In [4], a quantification of fuzziness in the measurement scale device is obtained. The natural follow-up question is how such fuzziness in data impacts the ensuing statistical inference, and if the above quantification can help in any way of making suitable adjustments in the statistical analysis. In the current work we pursue that study with specific focus to inference problems involving critical but basic parameters, such as population mean, variance and proportion. There is increasing interest in research addressing these questions and beyond, that is statistical analysis based on fuzzy data, [see e.g. [1], [3], [4], [5], [6], among others and the references therein], although the approaches vary greatly.

Thus, of interest is $\theta = \theta(\Im)$, a suitable parameter of $\Im$, the distribution of X. However X is not observable. Corresponding to n (unobservable) are i.i.d. values $X_1, X_2, \ldots, X_n$ from the population $\Im$, what we observe are $Y_1, Y_2, \ldots, Y_n$, which are i.i.d. observations, with Y_i being fuzzy form of X_i, as prompted by the measurement scale device. In our consideration, X_i's are typically continuously distributed unobservable quantities, while Y_i's are in a k-point Likert scale with different models /forms of fuzziness. The resultant fuzziness of scale has been quantified in [4] as:

$$\Psi = \frac{E[Var(X|Y)]}{Var(X)}. \tag{1}$$

The role of k in the above measureis de-emphasized here, although in the computational illustrations we experiment with few alternative values of k.

In the three inference problems of immediate interest, the parameter is

1. $\theta = \mu = \mathbb{E}(X)$;
2. $\theta = \sigma^2 = \text{Var }(X)$;
3. $\theta = \pi_c = \mathbb{P}(X_1 > X_2)$.

Of critical importance and focus is the comparison of parameters between two populations and study the extent by which relative values of the two characteristics (mean, proportion or variance of two variables/ population) remain unaltered in light of the fuzziness of data.

To illustrate, we take two practical examples — i) an academic institution taking students' feedback for courses and faculty ii) a hotel management collecting customer feedback on various kinds of services it provides. In the following we draw upon some queries in the context of students' feedback taken in a 5 point Likert scale for courses taught in multiple sections (same course by same/different faculty in different section) or different courses taught in the same section:

1. Is there significant evidence that a faculty (who taught both Sections A and B) performed worse in Section A as compared to Section B?

2. Is the variability of course rating the same between the two sections?

3. Is the true average rating received in a course at least 2.5 out of 5?

4. Do the majority of the students find the course workload to be reasonable?

Similar questions from the second contextual example pertaining to the service at reception, room service or drawing comparison between two competing restaurants in the Hotel compund.

2 Model of Fuzziness

We consider the following two structures or models of fuzziness.

1. Model 1: The crisp model under which the relationship between Y and X is unambiguous, even though unknown.

$$Y = i \Leftrightarrow \alpha_{i-1} < X \leq \alpha_i, \quad i = 1, \ldots, k$$

   where $\alpha_0 = -\infty$, $\alpha_{k+1} = \infty$ and the remaining α_i's are either known/ specified parameters in the model or are unknown parameters meant to be estimated from data.

2. Model 2: This is typical fuzzy-theoretic model under which,

$$P(Y = i|X = x] \propto \begin{cases} 1 & \text{if } (i = 1 \text{ and } x < 1) \text{ or } (i = k \text{ and } x > k) \\ t(\frac{x-i}{\beta}) & \text{otherwise} \end{cases};$$

   where $t(\cdot)$ is a suitably chosen function, e.g. $t(y) = \max(0, 1 - |y|)$, or $t(y)\mathbb{I}_{|y|<1}$ or $t(y) = exp(-t^2)$, and β is a model parameter to be prefixed or estimated.

Consider the following illustration of Model 1, where $X \curvearrowright N(\mu_x, \sigma_x^2)$. Let $a_i = \frac{\alpha_i - \mu_x}{\sigma_x}$. Then $p_i = P(Y = i) = \Phi(a_i) - \Phi(a_{i-1})$ and $E(Y) = \mu_y = \sum_{i=1}^{k} i \times p_i$, the conditional distribution of X given $Y = i$ is the (re-scaled) Normal distribution truncated between α_{i-1} and α_i, i.e. the part of the unconditional distribution, and hence

$$E(X|Y = i) = \mu_x + \sigma_x \times \frac{\phi(a_{i-1}) - \phi(a_i)}{p_i},$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ respectively denote the p.d.f. and C.D.F. of the standard normal distribution.

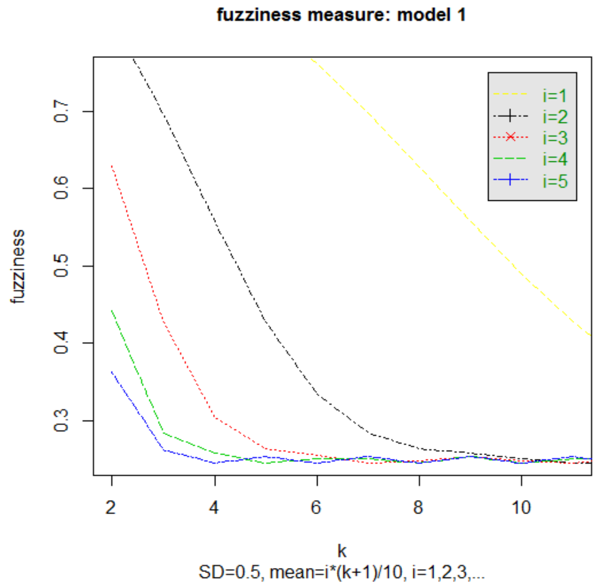
$$Var(X|Y = i) = \sigma_x^2 \times \left[1 + \frac{a_{i-1}\phi(a_{i-1}) - a_i\phi(a_i)}{p_i} - \frac{\{\phi(a_{i-1}) - \phi(a_i)\}^2}{p_i^2}\right]$$

and hence the fuzziness measure is given by

$$\begin{aligned} \frac{E[Var(X|Y)]}{Var(X)} &= \sum_{i=1}^{k} p_i \times \left[1 + \frac{a_{i-1}\phi(a_{i-1}) - a_i\phi(a_i)}{p_i} - \frac{\{\phi(a_{i-1}) - \phi(a_i)\}^2}{p_i^2}\right] \\ &= 1 - \sum_{i=1}^{k} \frac{\{\phi(a_{i-1}) - \phi(a_i)\}^2}{p_i} \end{aligned}$$

We may note that $\sigma_y^2 \neq Var[E(X|Y)]$ and indeed σ_y can be bigger than σ_x.

Figure 1: Fuzzy measure in Model 1 example



While the choice of scale is not the focus of this paper, Figure 1 suggests that unless the characteristics is of extreme nature, Likert scale with 5 options may suffice.

Special Model of fuzziness

For some model of fuzziness, it is possible that:

$$E(X|Y) \equiv Y. \tag{2}$$

This implies that the fuzziness of the scale would not have any primary (direct) impact on the basic inference for mean, since 2) implies $\mu_x = \mu_y$.

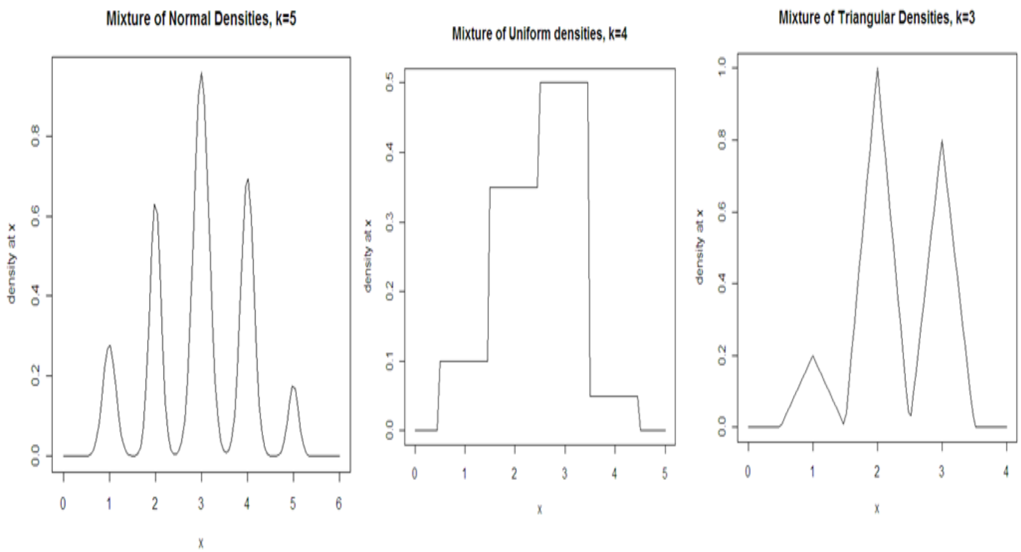
In the context of the above inference for mean, as well as in the case of independent focus on inference for variance, $\theta = \sigma^2(\Im)$, it is important to note the role of the measure of fuzziness, through the relation:

$$\sigma_Y^2 = \sigma_X^2(1 - \Psi) \tag{3}$$

We study the complete significance of (3), in the context of inference problems for both mean and variance.

Some examples where (2) is valid are different mixtures of symmetric distributions as model for X (with suitable cases of Model 1 and Model 2 are given in Figure 2.

Figure 2: Few Examples where Assumption 1 is valid



However (2) does not hold good in many situations/models, including the simple case of X following a Normal distribution. In such cases, we study the degree of difference in the two means, leading to validity of the approximation.

3 Inference of Mean and Variance

3.1 One Sample Mean Problem

We next look at how the mean of the observed (fuzzy) changes with the true mean under Model 1. In Figure 3a, we have plotted the observed mean (of Y) vs. the true mean (of X), while in Figure 3b, we have plotted the difference in means (between the observed and the true) vs. the true mean. In both the figures, plot is drawn for the different levels of variability in the true X, namely its standard deviation being 0.4, 0.6, 0.8 and 1.0.

Figure 3a: Observed vs. True Mean: Model 1 with 5 point Likert Scale

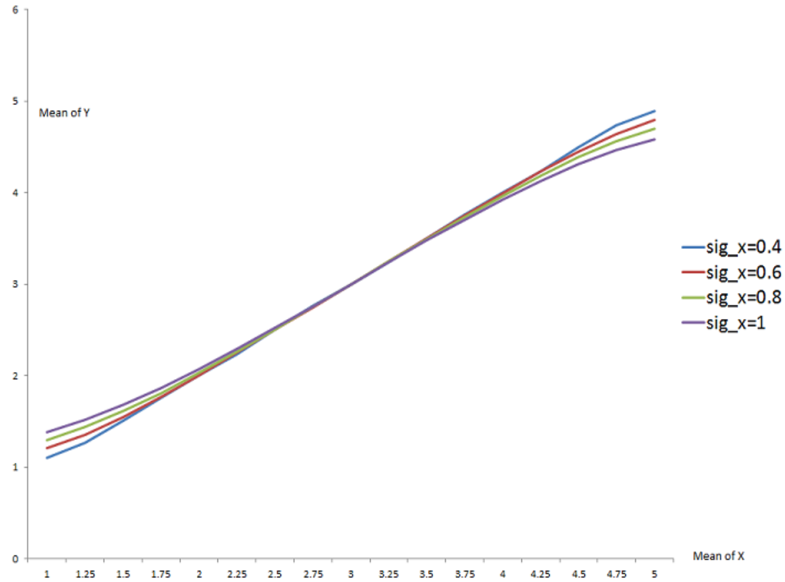
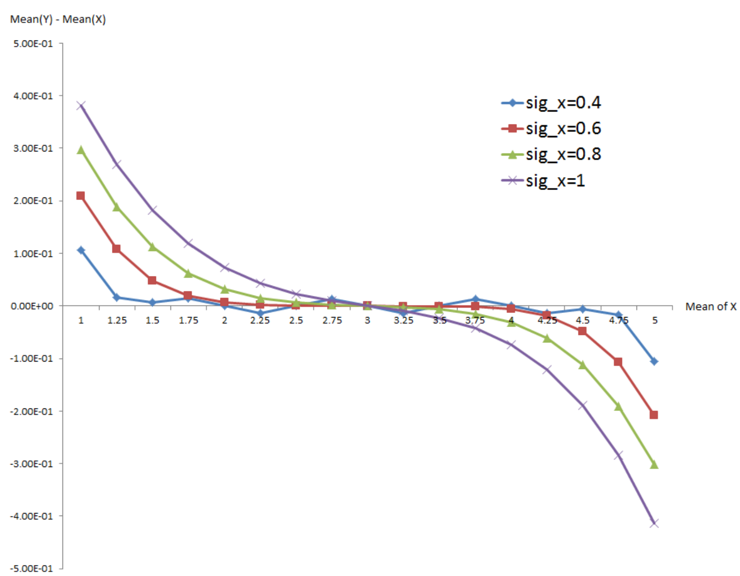


Figure 3b: Difference between Observed and True Mean: Model

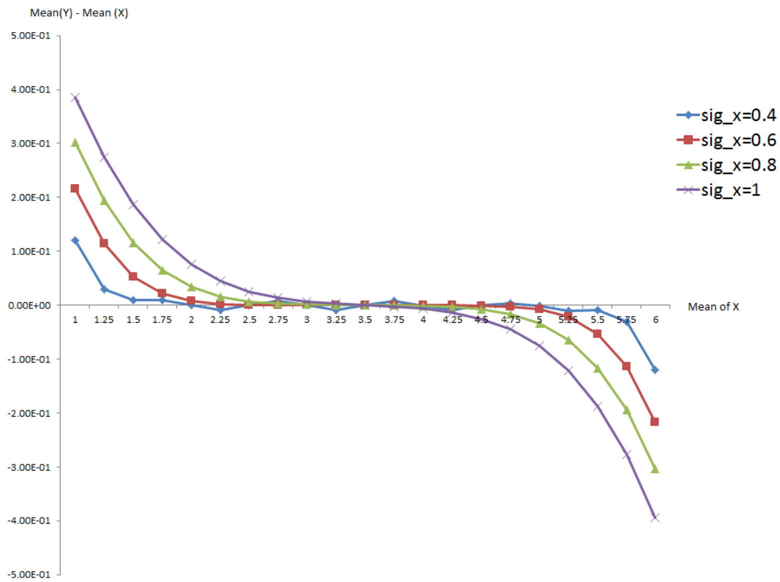


1

Considering other scales and Model 2, the above comparison remains essentially the same. For example, while considering Model 2 with $t(u) = \max(0, 1 - |u|)$, and the observed values being in 6 point Likert scale, we see the following comparison between the observed and the true means.

Figure 4: Difference between Observed and True Mean:

Model 2 with $t(u) = \max(0, 1 - |u|)$, and Y in 6 point Likert Scale



We summarize these by making the following observations with respect to inference of mean from fuzzy data:

- The *error* in mean is more for extreme characteristics.
- When the true mean is extreme (either low or high), the observed mean under-represents the extremities. That is, the mean from the fuzzy data is somewhat closer to the centre than the true mean.
- For broadly average characteristics, the difference between the observed (fuzzy) and the true mean is very nominal.
- The difference between the observed and the true mean increases with increase in true SD.
- The above difference is not very sensitive with respect to
  - the choice of scale (k), the number of options in the Likert scale;
  - the model of fuzziness, i.e. the crisp or fuzzy set approach in terms of Model 1 vis--vis Model 2;
  - the choice of membership function, when Model 2 is considered.

3.2 One Sample Variance Problem

Next we turn our attention to the comparison between the true and the observed fuzzy variance; we do this by looking at the ratio of the two under the two models considered above and sample diagrams are given below:

Figure 5a: Ratio of Observed to True Variance: Model 1 with $k = 5$

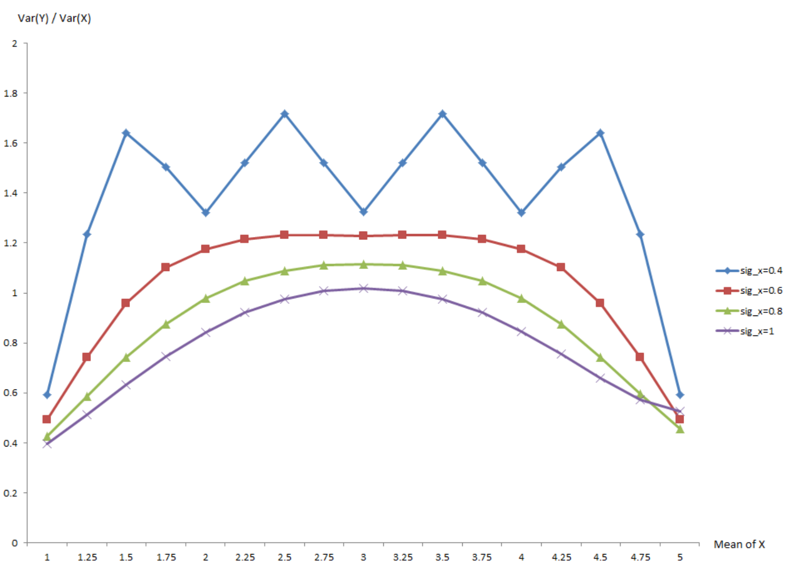
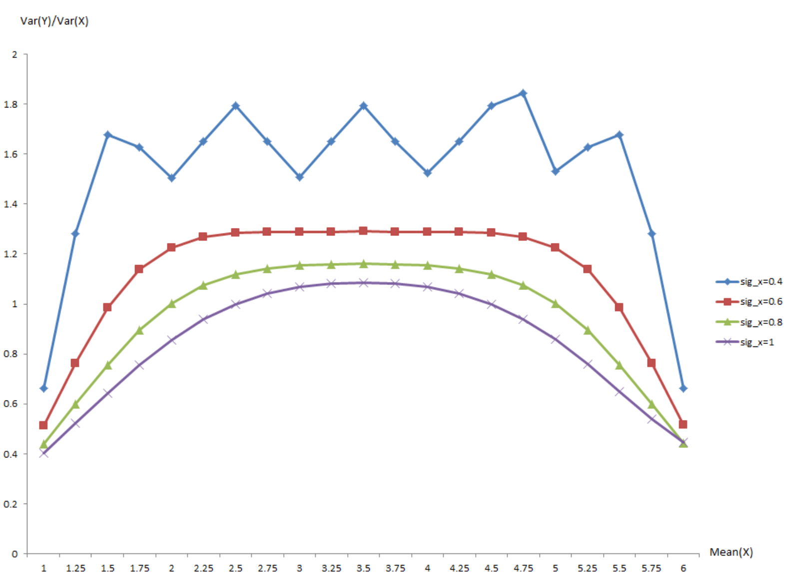


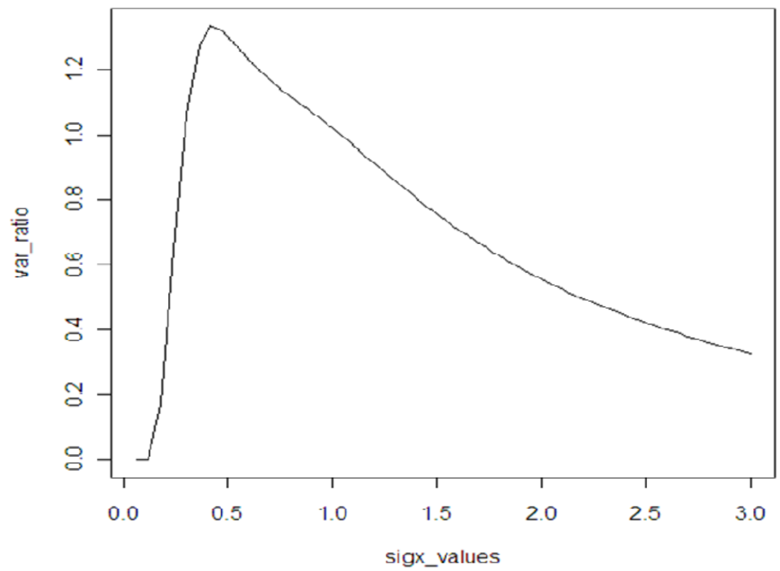
Figure 5b: Ratio of Observed to True Variance: Model 2 with $k = 6$



To probe further how the ratio of observed to true ratio changes with true standard deviation, we look at Figure 6.

Figure 6: Ratio of Observed to True Variance: how it changes with true SD

Model 1 with $k = 5$, true mean = 3



To summarize the observations for comparison of variance, we note:

- When the true standard deviation is below a threshold value, as per our Model 1, there is no variability due to fuzziness in the observed Y; even in Model 1, this is much lower. This explains the peculiarity in the observed Figures 5a and 5b (also confirmed in Figure 6) when the true variance is very low.

- Otherwise the observations on the error is similar to the case when the characteristic is mean in terms of being less for central characteristics and more for the extreme ones.

- Underestimation of the true variance because of fuzziness increase with increase in true variance.

- These observations appear not to be sensitive with respect to the choice of Model 1 or 2, or membership function or the number of options in the Likert scale.

4 Two Sample Inference

Next we turn our attention to what extent, the fuzziness of the data impacts comparison of mean between two populations. We assume the same model of fuzziness, as would be typically reasonable in the context given the same scale being used. Certain empirical observations can be extended from the one-sample study. For example, we empirically noted that given Model 1 and 7 point Likert scale being used, if $\sigma_X \leq 0.6$,

$$2 \leq \mu_X \leq 6 \quad \Rightarrow \quad |\mu_Y - \mu_X| \leq 0.008.$$

This can now be used to conclude that

$$\sigma_{X_i} =\leq 0.6, \text{and } 2 \leq \mu_{X_i} \leq 6,\ i = 1, 2 \quad \Rightarrow \quad ||\mu_{Y_1} - \mu_{Y_2}| - |\mu_{X_1} - \mu_{X_2}|| \leq 0.016.$$

This empirically suggests that comparison of mean based on fuzzy observations would be reasonable unless the attributes are very extreme.

Similarly, we have noted from one-sample study that unless the true variability is really small (or to some extent, it is really large), the ratio between observed and true variance is essentially constant with respect to change in mean. These observations from the one-sample study may be judiciously used to suggest that any comparison of variance, even for average attribute and having not very small volatility, inference from fuzzy observation may be misleading, but it would be possibly to correct it.

4.1 Two Sample Testing of Hypothesis of Mean

We now move on to more objective study and now consider Testing of hypothesis framework of 2 sample mean comparison. Thus, the objective is to answer questions like: "Is the rating in Section 1 higher on average compared to Section 2?", and inspect the consequence of inference from fuzzy data as opposed to (unobservable) ideal data in that context. For convenience and illustration, we consider the large sample situation only.

Thus the objective is to test $H_0 : \mu_{X_1} = \mu_{X_2} (= \mu)$ vs. $H_a : \mu_{X_1} > \mu_{X_2}$. However, the ideal observations $X_{i,j},\ i = 1, 2;\ j = 1, \ldots, n_i$, are not observed and instead if we observe the $Y_{i,j}$'s. Hence, while the ideal critical region at level α would have been

$$\frac{\bar{X}_1 - \bar{X}_2}{\frac{S^2_{x,1}}{n_1} + \frac{S^2_{x,2}}{n_2}} > z_\alpha$$

in practice, one may end up using:

$$\frac{\bar{Y}_1 - \bar{Y}_2}{\frac{S^2_{y,1}}{n_1} + \frac{S^2_{y,2}}{n_2}} > z_\alpha$$

where

$$\bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{i,j};\ \bar{Y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{i,j}; \quad i = 1, 2; \quad z_\alpha = \Phi^{-1}(1 - \alpha).$$

$$S^2_{x,i} = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (X_{i,j} - \bar{X}_i)^2;\ S^2_{y,i} = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (Y_{i,j} - \bar{Y}_i)^2;\ i = 1, 2.$$

As an illustration, we consider Model 1 with observations being 5 point Likert scale. Both the actual populations are Normal with standard deviation of 0.5. Sample sizes are respectively 60 and 65. Enclosed in Table 1 are comparison of actual (using fuzzy data Y's) and ideal (using non-fuzzy X's, which are unobservable) probability of committing type I error when α=0.05, using a simulation size of 100,000.

Table 1: Probability of Type I error with Ideal and fuzzy data in Model 1 example:

	μ								
	0.5	0.75	1	1.25	1.5	1.75	2	2.25	2.5
Ideal data	0.051	0.050	0.052	0.049	0.051	0.050	0.053	0.050	0.052
fuzzy data	0.051	0.049	0.052	0.050	0.051	0.052	0.053	0.051	0.051
	μ								
	2.75	3	3.25	3.5	3.75	4	4.25	4.5	
Ideal data	0.052	0.052	0.051	0.051	0.051	0.051	0.051	0.051	
fuzzy data	0.052	0.051	0.052	0.049	0.052	0.050	0.051	0.049	

Table 1 shows that there is hardly any impact of data fuzziness on probability of Type I error. Also the situation does not change with the true mean.

Next we look at the impact of fuzziness on the power of the above test. Table 2 shows that the test with fuzzy data is lot less powerful. Also what matters is the difference in μ, not the values themselves, as we know theoretically with ideal data.

Table 2: Comparison between Actual and Ideal Power in Model 1 example:
Power of the tests when $\mu_{X_2} - \mu_{X_1}$=0.25

	μ								
	0.5	0.75	1	1.25	1.5	1.75	2	2.25	2.5
Ideal data	0.874	0.876	0.875	0.873	0.874	0.876	0.876	0.874	0.875
fuzzy data	0.780	0.780	0.778	0.782	0.781	0.781	0.782	0.781	0.782
	μ								
	2.75	3	3.25	3.5	3.75	4	4.25	4.5	
Ideal data	0.874	0.875	0.872	0.873	0.873	0.874	0.875	0.872	
fuzzy data	0.778	0.779	0.779	0.782	0.778	0.778	0.781	0.782	

We attempt to probe deeper by bringing out how this efficiency, (power with fuzzy data as a percentage of power with ideal data), depends on difference of means via Table 3.

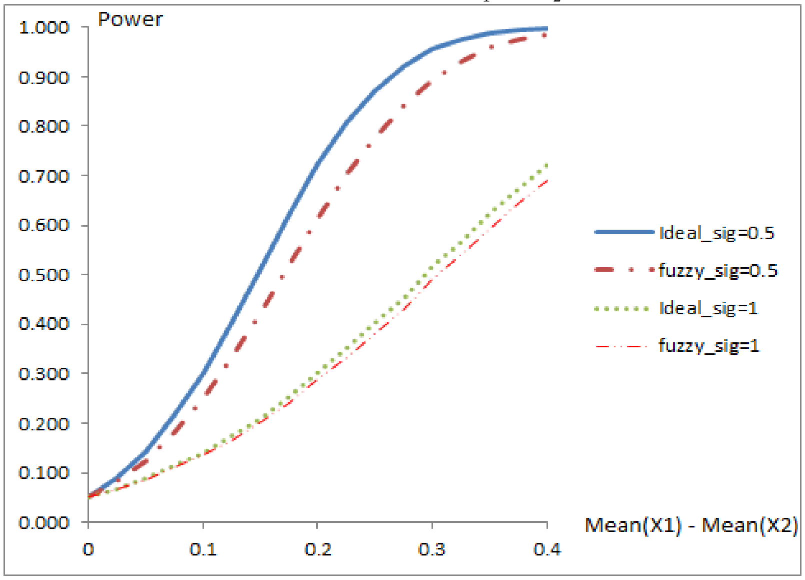
Diagrammatically this is represented in Figure 7, where we have two sets of tests corresponding to the power comparison between tests with ideal data and fuzzy, once when the true (common) standard deviation is 0.5, and the other when it is 1.0. This suggests that although the efficiency of the test is smaller with higher variability both with ideal as well as fuzzy data, the relative loss of efficiency with fuzzy data is somewhat smaller in the case of higher variability.

Table 3: Actual and Ideal Power with changing $\mu_{X_2} - \mu_{X_1}$: Equal Variance

μ_{X_1}	μ_{X_2}	Power with Ideal data	Power with fuzzy data	Efficiency Percentage
3	3	0.051	0.051	99.69%
3.025	3	0.087	0.082	94.07%
3.05	3	0.142	0.124	87.66%
3.075	3	0.215	0.181	84.45%
3.1	3	0.300	0.250	83.24%
3.125	3	0.401	0.332	82.68%
3.15	3	0.515	0.424	82.40%
3.175	3	0.621	0.519	83.48%
3.2	3	0.721	0.613	85.03%
3.225	3	0.808	0.702	86.92%
3.25	3	0.872	0.778	89.14%
3.275	3	0.921	0.842	91.45%
3.3	3	0.955	0.894	93.63%
3.325	3	0.977	0.932	95.35%
3.35	3	0.989	0.959	96.96%
3.375	3	0.994	0.976	98.15%
3.4	3	0.997	0.986	98.85%
3.5	3	1.000	0.999	99.92%

Figure 7: Power comparison of 2 sample Mean test:

Fuzzy data vs. Ideal data. $\sigma_{X_1} = \sigma_{X_2} =$ **0.5 or 1**



We then explore if relaxing equal standard deviation assumption changes the observations. Towards that we modify the above example by taking $\sigma_{X_1} = 0.5$, and $\sigma_{X_2} = 0.75$. The power comparisons are given in the following Table 4.

Table 4: Actual and Ideal Power with changing $\mu_{X_2} - \mu_{X_1}$: Unequal Variance

		Power with		Efficiency
μ_{X_1}	μ_{X_2}	Ideal data	fuzzy data	Percentage
3	3	0.051	0.051	99.71%
3.025	3	0.078	0.075	96.65%
3.05	3	0.119	0.110	92.49%
3.075	3	0.164	0.147	90.02%
3.1	3	0.228	0.201	88.31%
3.125	3	0.296	0.260	87.90%
3.15	3	0.376	0.330	87.59%
3.175	3	0.465	0.403	86.66%
3.2	3	0.548	0.479	87.52%
3.225	3	0.634	0.560	88.30%
3.25	3	0.714	0.638	89.29%
3.275	3	0.782	0.709	90.62%
3.3	3	0.842	0.773	91.86%
3.325	3	0.888	0.827	93.11%
3.35	3	0.925	0.874	94.56%
3.375	3	0.951	0.911	95.79%
3.4	3	0.970	0.939	96.85%

We conclude this subsection with a summary of observations in the two sample testing of hypothesis of mean. Fuzzy data does make much of an impact on the Type I error, at least in the large sample case. However, the power of the test is reduced with fuzzy data. Higher variability leads to less power, however the inference from fuzzy data is closer to the case with full data. In this context, it does not matter so much whether the true variability of the variables being compared are equal or not. There is a need further extension of the conclusions in light of changing sample size and robustness of the population.

4.2 Two Sample Variance Problem

Next we move on to two sample Variance Comparison. That is, we attempt to look at impact of fuzziness in responses to queries like: "Is the variability of rating in Section 1 significantly higher as compared to Section 2?"

The framework is similar to that in the two-sample variance problem — we now focus on testing

$H_0 : \sigma_{X_1} = \sigma_{X_2}$ vs. $H_a : \sigma_{X_1} > \sigma_{X_2}$. The critical region for a test of size α with ideal data is:

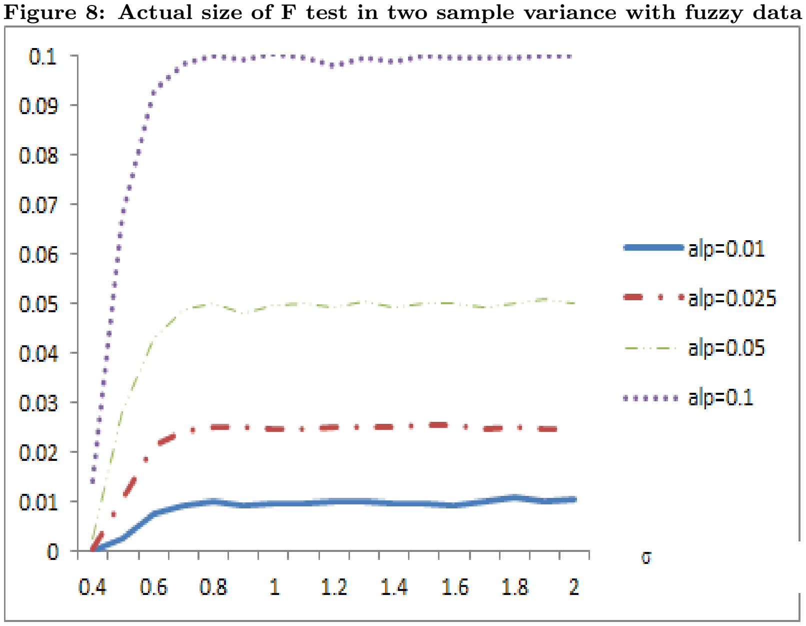
$$\frac{S^2_{x,1}}{S^2_{x,2}} > f_{n_1-1,n_2-1,\alpha},$$

while with fuzzy data, we have to use:

$$\frac{S^2_{y,1}}{S^2_{y,2}} > f_{n_1-1,n_2-1,\alpha}.$$

The size or actual probability of committing Type I error of the F-tests with fuzzy data for the intended tests of size α=0.01, 0.025, 0.05 and 0.1 is depicted in Figure 8 for varying values of σ ($= \sigma_{X_1} = \sigma_{X_2}$). [With ideal data, naturally, there is no difference between intended and actual size.]

Figure 8: Actual size of F test in two sample variance with fuzzy data



Next we compare the power of test using fuzzy data vis-a-vis ideal data in Table 5a and 5b.

Table 5a: Power of the F test in two sample variance TOH with ideal data

α	$k = \frac{\sigma_{X_1}}{\sigma_{X_2}}$									
	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
0.01	0.056	0.181	0.388	0.615	0.795	0.907	0.964	0.988	0.996	0.999
0.025	0.112	0.294	0.532	0.746	0.885	0.955	0.984	0.995	0.999	1.000
0.05	0.185	0.413	0.655	0.835	0.933	0.978	0.993	0.998	1.000	1.000
0.1	0.294	0.558	0.776	0.910	0.969	0.991	0.998	0.999	1.000	1.000

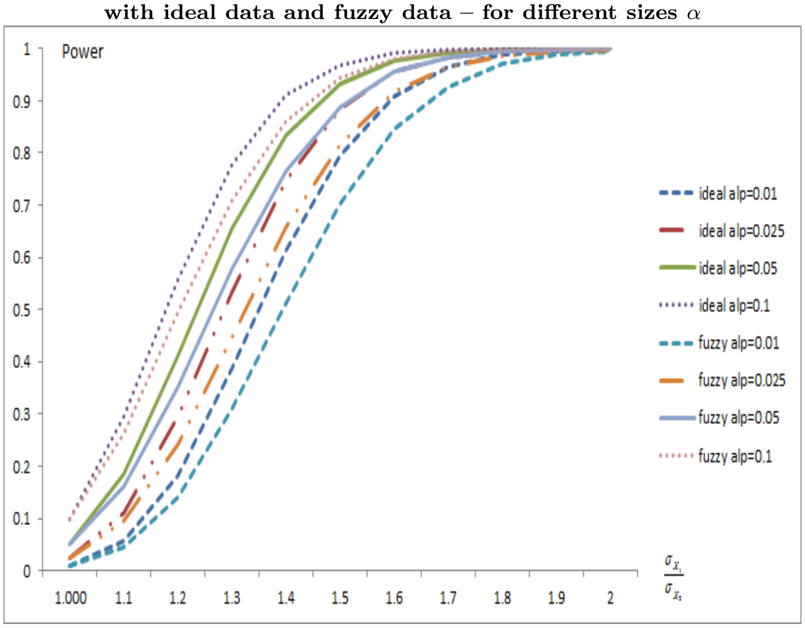
Table 5b: Power of the F test in two sample variance TOH with fuzzy data

α	$k = \frac{\sigma_{X_1}}{\sigma_{X_2}}$									
	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
0.01	0.046	0.141	0.310	0.512	0.704	0.846	0.927	0.970	0.988	0.996
0.025	0.095	0.240	0.447	0.659	0.818	0.917	0.966	0.987	0.996	0.999
0.05	0.162	0.351	0.577	0.765	0.889	0.956	0.984	0.995	0.999	0.999
0.1	0.263	0.494	0.708	0.862	0.944	0.979	0.993	0.998	0.999	1.000

In Figure 9, we present comparison of powers, between tests with ideal data and fuzzy data, for four sizes, namely with the Type of error α being 1%, 2.5%, 5%, and 10%.

Figure 9: Power of F tests in two sample variance TOH:

with ideal data and fuzzy data – for different sizes α



To summarize the case of testing of hypothesis in the two sample variance problem, we note that the Type I error using fuzzy data is grossly under-represented, if the true variabilities are very small. Of course, this is intuitively as well as theoretically justified. However, the inference is reliable, if true variabilities are large (not so small). Power of the test using fuzzy data is less, with the difference getting accentuated with relative increase in alternate variability, before finally the two functions eventually merge (naturally, as they have to).

4.3 Paired Comparison of Proportion: Sign Test of Proportion

In this section, we inspect the impact of fuzziness on the sign test — namely, the paired comparison of proportion. So for example, the relevant queries are: "Is the proportion of time Instructor 1 getting higher rating than Instruction 2 significantly higher than the other way round?" The rating (as usual in the Likert scale) are given by the same set of students, which is the basis of the paired comparison. Thus, we wish to test

$$H_0 : P(X_1 > X_2) = P(X_1 < X_2) \quad \text{vs.} \; H_a : P(X_1 > X_2) > P(X_1 < X_2).$$

The random variables are paired, i.e. the observations are $(Y_{1,i}, Y_{2,i}),\ i = 1, \ldots, n$ which are fuzzy versions of ideal observations $(X_{1,i}, X_{2,i}),\ i = 1, \ldots, n$. Let f_x^1 and f_y^1 respectively denote the number of X-pairs and Y-pairs, where the first element is higher, and similarly we define f_x^2 and f_y^2 i.e.

$$f_x^1 = \sum_{i=1}^{n} \mathbb{I}_{\{X_{1,i} > X_{2,i}\}};\ f_x^2 = \sum_{i=1}^{n} \mathbb{I}_{\{X_{1,i} < X_{2,i}\}}; \quad f_y^1 = \sum_{i=1}^{n} \mathbb{I}_{\{Y_{1,i} > Y_{2,i}\}};\ f_y^2 = \sum_{i=1}^{n} \mathbb{I}_{\{Y_{1,i} < Y_{2,i}\}}.$$

We also denote:

$$n_x = f_x^1 + f_x^2, \quad p_x = \frac{f_x^1}{n_x}; \qquad n_y = f_y^1 + f_y^2, \quad p_y = \frac{f_y^1}{n_y};$$

noting that n_x and n_y can be different (and less than n) because of ties. For large n (consequently large n_x or n_y) The critical region for a test of size α with ideal data is:

$$\frac{p_x - 0.5}{\sqrt{\frac{0.5 \times 0.5}{n_x}}} = \sqrt{n_x}(2p_x - 1) > z_{1-\alpha},$$

while with fuzzy data, we have to use:

$$\sqrt{n_y}(2p_y - 1) > z_{1-\alpha}.$$

First we look at impact of fuzziness on the size of test. For initial exploration, we consider Model 1 with 5-point Likert scale observations while the true standard deviations are 0.5 and sample size is 75. The comparison in terms of size (α) of the test is given in Table 6a below:

Table 6a: Comparison of size of Paired tests: Fuzzy data with ideal data

	α=0.05		α=0.01	
μ	ideal data	fuzzy data	ideal data	fuzzy data
0.5	0.051	0.050	0.010	0.009
1.5	0.053	0.051	0.009	0.010
2.5	0.053	0.049	0.010	0.010
3.5	0.052	0.049	0.010	0.010
4.5	0.054	0.049	0.010	0.010

Since the initial simulation suggests almost no less efficiency using the fuzzy data and in fact for $\alpha = 0.05$, it seems to be working better than even ideal data, we explore in greater depth, including expanding the scope by changing standard deviations and considering two sample sizes, names 75 and 200. The results are given in Tables 6b and 6c.

Table 6b: Comparison of size of Paired tests: sample size =75

	P(type I error with actual data X's)				P(type I error with fuzzy data Y's)			
$n = 75$	$\sigma_{X_1} = \sigma_{X_2} = \sigma$				$\sigma_{X_1} = \sigma_{X_2} = \sigma$			
α	0.4	0.6	0.8	1	0.4	0.6	0.8	1
0.01	0.010	0.010	0.010	0.010	0.009	0.010	0.010	0.010
0.025	0.032	0.032	0.034	0.033	0.026	0.024	0.026	0.025
0.05	0.053	0.053	0.052	0.053	0.049	0.048	0.049	0.050
0.1	0.082	0.082	0.082	0.084	0.098	0.101	0.103	0.101

Table 6c: Comparison of size of Paired tests: sample size =200

	P(type I error with actual data X's)				P(type I error with fuzzy data Y's)			
$n = 200$	$\sigma_{X_1} = \sigma_{X_2} = \sigma$				$\sigma_{X_1} = \sigma_{X_2} = \sigma$			
α	0.4	0.6	0.8	1	0.4	0.6	0.8	1
0.01	0.009	0.010	0.009	0.009	0.010	0.010	0.009	0.010
0.025	0.029	0.028	0.028	0.028	0.025	0.024	0.024	0.025
0.05	0.053	0.051	0.052	0.051	0.050	0.049	0.050	0.048
0.1	0.091	0.089	0.090	0.091	0.102	0.100	0.099	0.101

Next we compare the power of the paired test through Table 7.

Table 7: Comparison of power of the Paired test: Fuzzy data vs. Ideal data

Diff. in mean	ideal data	fuzzy data	percentage
0	0.052	0.051	98.23%
0.025	0.086	0.083	96.12%
0.05	0.128	0.126	98.50%
0.075	0.188	0.185	98.48%
0.1	0.263	0.261	99.07%
0.125	0.346	0.350	101.17%
0.15	0.441	0.447	101.43%
0.175	0.537	0.548	102.07%
0.2	0.629	0.647	102.77%
0.225	0.714	0.734	102.82%
0.25	0.792	0.810	102.34%
0.275	0.854	0.869	101.83%
0.3	0.901	0.916	101.56%
0.325	0.939	0.951	101.25%
0.35	0.962	0.970	100.83%
0.375	0.978	0.983	100.56%
0.4	0.988	0.991	100.28%

In summary, fuzziness does not hurt the paired Sign test. The observed (actual) Type I error, as obtained from fuzzy data is as good as that from the ideal data. Even the power of the test with fuzzy data is comparable with that with ideal data or even sometimes better. Of course a much more extensive study is in order to investigate this finding which is somewhat puzzling.

5 Future Directions

In this preliminary work, the accuracies in basic statistical inferences involving mean, variances and proportion, for single and two populations, have been investigated under few models and also on the basis of certain simulation studies. Naturally there is a need and scope to broaden the framework before more general conclusions can be arrived. Also it is necessary to move beyond not only to multiple populations studies, e.g. in ANOVA, but also to multivariate analysis like principal component analysis, cluster analysis etc. these are planned to be taken up in future.

References

[1] Statistical Modeling, Analysis and Management of Fuzzy Data, Edited by Bertoluzza, Carlo; Gil, Maria A.; Ralescu, Dan A. Series: Studies in Fuzziness and Soft Computing, Vol. 87, 2002, Physica Verlag Heidelberg.

[2] A.C. Boyce, Methods of measuring teachers' efficiency, Fourteenth Yearbook Nat. Soc. Stud. Educ., Pt II, 1915.

[3] R. Coppia, M.A. Gilb, H.A.L. Kiersc . The fuzzy approach to statistical analysis, Computational Statistics & Data Analysis, 51 (1), 2006, 1-14.

[4] S. Das, Quantifying fuzziness due to the scale of measurement in response systems, Fuzzy Sets and Systems 132 2002: pp 317 - 333.

[5] S. Frhwirth-Schnatter, On statistical inference for fuzzy data with applications to descriptive statistics, Fuzzy Sets and Systems 50 , 1992, 143-165.

[6] R. Viertl, Statistical Methods for Fuzzy Data (Wiley Series in Probability and Statistics), 2010.