Hypotheses testing as a fuzzy set estimation problem

Glen Meeden*

School of Statistics

University of Minnesota

Minneapolis, MN 55455

glen@stat.umn.edu

Siamak Noorbaloochi

Center for Chronic Disease Outcomes Research

Minneapolis VA Medical Center

Minneapolis, MN 55417

and Department of Medicine

University of Minnesota

Siamak.Noorbaloochi@med.va.gov

July 08

*Research supported in part by NSF Grant DMS 0406169

SUMMARY

For many scientific experiments computing a p-value is the standard method for reporting the outcome. Once the hypotheses testing problem has been formulated it yields a simple way of summarizing the information in the data. One theoretical justification for p-values is the Neyman-Pearson theory of hypotheses testing. However the decision making focus of this theory does not correspond well with the desire, in most scientific experiments, for a simple and easily interpretable summary of the data. Fuzzy set theory with its notion of a fuzzy membership function gives a non-probabilistic way to talk about uncertainty. Here we argue that for some situations where a p-value is usually computed it may make more sense to formulate the question as one of estimating a fuzzy membership function. This function will be the fuzzy membership function of the subset of special parameter points which are of particular interest for the experiment. Choosing the appropriate fuzzy membership function can be more difficult than specifying the null and alternative hypotheses but the resulting payoff is greater. This is because a fuzzy membership function can better represent the shades of desirability among the parameter points than the sharp division of the parameter space into the null and alternative hypotheses. The new approach yields an estimate which is easy to interpret and more flexible and informative then the cruder p-value.

AMS 1991 subject classifications Primary 62F03; secondary 62G05.

Key Words and phrases: Fuzzy set theory; membership function; p-values; hypotheses testing; point estimation.

1 Introduction

The concept of p-value or level of significance is widely used in practice to measure the strength of evidence against a null hypothesis. The usual formal justification comes from the Neyman-Pearson theory of hypotheses testing which assumes a sharp break between the null and alternative hypotheses and the necessity of making an accept or reject decision. Both of these assumptions make little sense in most scientific work where a simple summary of the information contained in the outcome of an experiment is desired. One approach which attempts to overcome some of these problems is the theory of equivalence testing which is discussed in Wellek (2003). Another modification of the theory allows for an indifference zone between the two hypotheses but this is little used in practice. Both of these alternatives highlight the fact that in many situations the choice of null and alternative hypotheses is not so straightforward. Another problem with the standard theory is that if the true state of nature is in the alternative but close to the boundary and the sample size is large then there is high probability the outcome will be statistically significant although most observers would agree that the result is of no practical importance.

Fuzzy set theory was introduced in Zadeh (1965) and is another approach to representing uncertainty. A fuzzy set A is characterized by its membership function. This is a function whose range is contained in the unit interval. At a point the value of the membership function is a measure of how much we think the point belongs to the set A. A fuzzy set whose membership function only takes on the values zero or one is called *crisp*. For a crisp set, the membership function is just the usual indicator function of the set.

For the most part statisticians have shown little enthusiasm for using this new terminology to describe uncertainty. In the 1970's Max Woodbury developed the notion of Grade of Membership for applications in the health sciences. This notion measures the degree of partial membership of an individual belonging to several possible classes. The theory is developed in some detail in Manton et al. (1994). Taheri (2003) gives a review of applications of fuzzy set theory concepts to statistical methodology. Casals et al. (1986) consider the problem of testing hypotheses when the data is fuzzy and the hypotheses are crisp and Filzmoser and Viertl (2004) introduced the notion of fuzzy p-values for such problems. Arnold (1996) and Taheri and Behboodian (1999) consider problems where the hypotheses are fuzzy and the data are crisp. Blyth and Staudte (1995) proposed a theory which stayed within the general Neyman-Pearson framework and provided a measure of evidence for the alternative hypothesis rather than an accept-reject decision. Dollinger et al. (1996) noted that this approach can be reformulated using fuzzy terminology. Singpurwalla and Booker (2004) have proposed a model which incorporates fuzzy membership functions into a subjective Bayesian setup. However they do not give them a probabilistic interpretation. Geyer and Meeden (2005) assumed that both the hypotheses and data are crisp and introduced the notion of fuzzy p-values and fuzzy confidence intervals.

Here we will argue that many scientific problems where a p-value is computed can be reformulated as the problem of estimating the fuzzy membership function of the set of good or useful or interesting parameter points. Rather than specifying a null and alternative hypotheses we will choose a fuzzy membership function to represent what is of interest in the problem at hand. We will see that the usual p-value can be interpreted as estimating a particular membership function. We believe this suggests that more attention should be paid to the fuzzy membership function being estimated. A more careful choice of this membership function will allow a better representation of the realities of the problem under consideration and will avoid some of the difficulties associated with standard methods.

2 Fuzzy set theory

We will only use some of the basic concepts and terminology of fuzzy set theory, which can be found in the most elementary of introductions to the subject (Klir and St. Clair, 1997).

A *fuzzy set* A in a space Θ is characterized by its *membership function*, which is a map $I_A : \Theta \to [0, 1]$. The value $I_A(\theta)$ is the "degree of membership" of the point θ in the fuzzy set A or the "degree of compatibility ... with the concept represented by the fuzzy set". See page 75 of (Klir, St. Clair, and Yuan, 1997). The idea is that we are uncertain about whether θ is in or out of the set A. The value $I_A(\theta)$ represents how much we think θ is in the fuzzy set A. The closer $I_A(\theta)$ is to 1.0, the more we think θ is in A. The closer $I_A(\theta)$ is to 0.0, the more we think θ is not in A.

A natural inclination for statisticians not familiar with fuzzy set theory is to try to give a fuzzy membership function a probabilistic interpretation. To help overcome this difficulty consider the following situation. You need to buy a car. Let Θ be the set of all possible cars for sale in your area. Let A be the fuzzy set of cars that you could consider owning. For each car in the area you can imagine assigning it a value between 0 and 1 which would represent the degree of membership of this particular car in the fuzzy set A. For a given car this depends on its age, condition, style, price and so forth. Here the fuzzy membership function measures the overall attractiveness of a car to you. After checking out several cars and assessing their fuzzy membership you will eventually select one to buy.

We will consider some standard problems where p-values are often computed. In each case we will discuss how a fuzzy membership function can be selected which realisticly captures much of the background of the problem. We will then discuss how the resulting function can be estimated using standard methods.

3 Normal distribution testing problems

3.1 An example

An important responsibility of the Veterans Administration (VA) is to monitor the health of veterans. The American Heart Association has made the following recommendations for the level of total blood cholesterol.

- Desirable: Less than 200 mg/dL.
- High risk: More than 240 mg/dL.
- Borderline high risk: Between 200-239 mg/dL.

The VA is interested in the mean cholesterol level of a cohort of coronary heart disease patients. They plan to take a random sample of individuals and observe their cholesterol levels. How should they analyze the resulting data assuming that they are sampling from a normal population with unknown mean θ and unknown variance σ^2?

One possibility is to compute a simple point estimate for θ and make an "informal" judgment about the status of the population. In practice this judgment depends not only on the value of θ but on the value of σ^2 as well. For example their attitude could be quite different for a population with $\theta = 220$ and $\sigma = 20$ than for one with $\theta = 220$ and $\sigma = 40$

A second possibility would be to calculate the p-value for testing $H : \theta \leq \theta_0$ against $K : \theta > \theta_0$ where θ_0 is some value to be determined. In this example it is not so clear how to choose θ_0. Moreover whatever value of θ_0 is selected it is wrong to think of it as a sharp cut point between good and bad values of the population mean. Furthermore the size of the resulting p-value and its interpretation will very much depend on this choice.

One way to more formally bring these concerns into an analysis is to use fuzzy set theory. To this end we let H denote the set of good parameter points where the cholesterol level of the population is of lesser concern. This is done by defining I_H, the fuzzy membership function of H, for each point in the parameter space. This sounds like a daunting task but we shall see that convenient families of possible fuzzy membership functions make it feasible. Furthermore the information used to select a specific fuzzy membership function for a given problem incorporates just the kind of things that are used to interpret a p-value sensibly. An advantage of this approach is that it helps to make such considerations more explicit in the inferential process.

In what follows we will discuss how to select a reasonable fuzzy membership function for this problem and discuss how the resulting function can be estimated.

3.2 One sided alternative with known variance

To introduce our ideas we begin with a one sided normal mean testing problem with know population variance. Then we will consider the more realistic problem with unknown variance. Let $X_1, X_2, \ldots, X_n$ be iid normal(θ, σ^2) where $\theta \in (-\infty, \infty)$ is unknown and σ^2 is known. Consider the testing problem

$$H : \theta \leq \theta_0 \quad \text{against} \quad K : \theta > \theta_0 \tag{1}$$

For a fixed $0 < \alpha < 1$ let $\phi(X, \alpha, \theta_0)$ be the UMP level α test of H against K. Recall

$$E_\theta(\phi(X, \alpha, \theta)) = \alpha \qquad \text{for } 0 \leq \alpha \leq 1 \text{ and } \theta \in \Theta \tag{2}$$

Let $I_H(\theta)$ be the indicator function of the hypothesis H. Then among all tests ϕ with $E_{\theta_0}\phi(X) = 0.5$ the test $\phi(X, 0.5, \theta_0)$ minimizes uniformly in θ

$$d(\theta) = \begin{cases} E_\theta(1 - \phi(X)) - I_H(\theta) & \text{for } \theta < \theta_0, \\ I_H(\theta) - E_\theta(1 - \phi(X)) & \text{for } \theta > \theta_0. \end{cases} \tag{3}$$

The Neyman-Pearson theory of hypotheses testing can be formulated as a special case of decision theory. Here it will be useful to think of it as a special kind of estimation problem where the function to be estimated is $I_H(\theta)$. Equation 3 shows that among all tests with $E_{\theta_0}\phi(X) = 0.5$ the estimator $1 - \phi(X, 0.5, \theta_0)$, as an estimator of $I_H(\theta)$, uniformly minimizes the bias. As the sample size n increases its expectation becomes steeper and steeper in the neighborhood of θ_0 and a better and better approximation of $I_H(\theta)$.

Let $P(X)$ denote the p-value coming from the UMP family of tests. If θ_0 is true then the distribution of $P(X)$ is uniformly distributed on the unit interval and $E_{\theta_0} P(X) = 0.5$. $P(X)$ is essentially a smoother version of $1 - \phi(X, 0.5, \theta_0)$ and its expectation behaves very much like $1 - E_\theta \phi(X, 0.5, \theta_0)$ for large n. Hence the usual p-value can be interpreted as an unbiased estimator of a function which is an approximation of $I_H(\theta)$. This suggests that instead of insisting considering only crisp hypotheses, one could consider an hypothesis described by a fuzzy membership function. In this formulation one specifies a fuzzy membership function, $I_H(\theta)$, which captures the vagueness in the specification of the break point θ_0 and which also can be estimated unbiasedly. For each θ this function measures how strongly we believe that θ belongs to the set of good parameter values. In the VA example the observed p-value is an estimate of the degree of membership of the population parameters belonging to the set of values where the population cholesterol is of little concern.

To reiterate, for us, the usual p-value is an unbiased estimator of its expectation. This expectation can be in interpreted as the fuzzy membership function of the set of good parameter values. This new interpretation highlights one of the problems with the standard p-value. What it is estimating depends on the sample size n. This is what causes it to be very significant with high probability for large values of n and values of θ where $\theta - \theta_0$ is positive but very small. As we remarked earlier this is an unappealing property.

If our suggestion is accepted then an alternate program comes to mind. Rather than estimating a fuzzy membership function which depends only on

θ_0 and n one should select the fuzzy membership to be estimated which more accurately reflects the facets of the problem at hand. The degree of membership of θ in the fuzzy set of good parameter points can depend on many factors. In a given problem a thoughtful assessment should yield a more sensible fuzzy membership function than the one estimated by the usual p-value. Instead of finding an unbiased estimator of a fuzzy membership function one could find its Bayes estimate for a given prior distribution. Note however that the choice of a prior distribution use different kinds of prior information than what is used to select a sensible fuzzy membership function. There is nothing Bayesian in the selection of the fuzzy membership problem.

3.3 A family of fuzzy membership functions

Let Φ denote the distribution function of the standard normal distribution. Then for $\lambda > 0$ we claim that the family of functions of the form

$$\Phi\Big(\frac{\lambda\sqrt{n}}{\sqrt{1+\lambda^2}}\frac{\theta_0-\theta}{\sigma}\Big) \tag{4}$$

gives a sensible class of possible fuzzy membership functions to replace the testing problem of equation 1. Before justifying this statement we find the best unbiased estimator for a member of this family.

Lemma 1. *Let $\bar{X} = \sum_{i=1}^n X_i/n$ and Φ be the distribution function of the standard normal distribution. Then*

$$E_\theta \Phi\Big(\lambda\sqrt{n}\frac{\theta_0-\bar{X}}{\sigma}\Big) = \Phi\Big(\frac{\lambda\sqrt{n}}{\sqrt{1+\lambda^2}}\frac{\theta_0-\theta}{\sigma}\Big)$$

Proof. Let
$$a = \frac{\theta - \theta_0}{\sigma/\sqrt{n}}$$

Then by the change of variable formula we have

$$\begin{aligned} E_\theta \Phi\Big(\lambda \frac{\theta_0 - \bar{X}}{\sigma/\sqrt{n}}\Big) &= \int_{-\infty}^{\infty} \Phi\Big(\lambda \frac{\theta_0 - \bar{x}}{\sigma/\sqrt{n}}\Big) \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp - \frac{(\theta - \bar{x})^2}{2\sigma^2/n}\, d\bar{x} \\ &= \int_{-\infty}^{\infty} \Phi(\lambda y) \frac{1}{2\pi} \exp - \frac{(y-a)^2}{2}\, dy \\ &= P(Z - \lambda Y \le 0) \end{aligned}$$

where Z and Y are independent and Z has the standard normal distribution and Y is normal$(a, 1)$. The result follows easily. □

We note in passing that if we let $\lambda = 1/\sqrt{n-1}$ then the function of θ in expression 4 becomes $P_\theta(X_1 \le \theta_0)$ and its estimator given in the lemma is its well known unbiased estimator. (Lehmann, 1986)

We can also use the lemma to find the expected value of the usual p-value for this problem. Let

$$\begin{aligned} p_v(\bar{x}) &= P_{\theta_0}(\bar{X} \ge \bar{x}) \\ &= 1 - \Phi\Big(\frac{\bar{x} - \theta_0}{\sigma/\sqrt{n}}\Big) \\ &= \Phi\Big(\sqrt{n}\frac{\theta_0 - \bar{x}}{\sigma}\Big) \end{aligned}$$

then for $\theta > \theta_0$

$$E_\theta p_v(\bar{X}) = E_\theta \Phi\Big(\sqrt{n}\frac{\theta_0 - \bar{X}}{\sigma}\Big) = \Phi\Big(\frac{\sqrt{n}}{\sqrt{2}} \frac{\theta_0 - \theta}{\sigma}\Big)$$

As n increase for $\theta \ne \theta_0$ this converges to the indicator function of the null hypothesis.

When trying to select a specific member of this family of functions for a particular problem the job becomes a bit easier if we let $a = \sqrt{n}\lambda/\sqrt{1+\lambda^2}$ and rewrite the equation in lemma 1 as

$$E_\theta \Phi\Big(\frac{a}{\sqrt{1-a^2/n}}\frac{\theta_0 - \bar{X}}{\sigma}\Big) = \Phi\Big(a\frac{\theta_0 - \theta}{\sigma}\Big) \tag{5}$$

The first step in selecting a fuzzy membership function is choosing the value θ_0. It will play a similar role to θ_0 in equation 1. Note for any choice of θ_0 every member of this family has the value 0.5 at θ_0. Hence θ_0 will be a "soft break" point between the good values of θ and the rest of the parameter space. The choice of 0.5 to be the value of our fuzzy membership function at θ_0 is somewhat arbitrary since it is really a question of calibrating the values of our function. We have chosen 0.5 because it agrees with the usual p-value at that point. It also gives the maximum range of possible values on either side of θ_0.

Once θ_0 has been selected it remains to choose a value for a. For a given a with $0 < a < \sqrt{n}$ the righthand side of the previous equation is easy to plot and an appropriate function could be selected by inspection. Alternatively one can choose a $\theta_1 > \theta_0$ and $0 < \beta < 0.5$ where β is the value of the fuzzy membership function at θ_1 and find the value of a which satisfies

$$\Phi\Big(a\frac{\theta_0 - \theta_1}{\sigma}\Big) = \beta \tag{6}$$

If z_β satisfies $\Phi(z_\beta) = \beta$ then clearly

$$a = \frac{-z_\beta \sigma}{\theta_1 - \theta_0} \tag{7}$$

where a must belong to the interval $(0, \sqrt{n})$. If not there is no solution.

An interesting special case of equation 4 occurs when we let $\lambda = \sigma/(\sqrt{n}\tau)$. In this case it becomes

$$E_\theta \Phi\Big(\frac{\theta_0 - \bar{X}_n}{\tau}\Big) = \Phi\Big(\frac{\theta_0 - \theta}{\sqrt{\sigma^2/n + \tau^2}}\Big) \tag{8}$$

In a particular case the value of τ can be determined by selecting θ_1 and $0 < \beta_1 < 1$ and using the equation

$$\Phi\Big(\frac{\theta_0 - \theta_1}{\sqrt{\sigma^2/n + \tau^2}}\Big) = \beta_1 \tag{9}$$

to solve for τ.

3.4 One sided alternative with unknown variance

Now we will consider the testing problem of equation 1 when the population variance is unknown. We assume the fuzzy membership function we wish to estimate is of the form

$$\Phi\Big(a\frac{\theta_0 - \theta}{\sigma}\Big) \tag{10}$$

where $a > 0$. The function depends on how far θ is from θ_0 in standardized units, i.e. corrected for the standard deviation. The choice of a controls how important a given standardized distance is in the fuzzy membership function.

We do not know an unbiased estimator for the function in equation 10. But we will find an approximate unbiased estimator that works very well. To that end let $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2/(n-1)$. Then

$$\begin{aligned} E_{\theta,\sigma}\Phi\Big(\lambda\frac{\theta_0 - \bar{X}}{S/\sqrt{n}}\Big) &= E_{\theta,\sigma}\Phi\Big(\frac{\lambda(\theta_0 - \theta)}{S/\sqrt{n}} - \frac{\lambda(\bar{X} - \theta)}{S/\sqrt{n}}\Big) \\ &= E_{\theta,\sigma}\Phi\Big(\frac{\sqrt{n}}{\sigma}\frac{\lambda(\theta_0 - \theta)}{\sqrt{S^2/\sigma^2}} - \lambda\frac{(\bar{X} - \theta)/(\sigma/\sqrt{n})}{\sqrt{S^2/\sigma^2}}\Big) \\ &= E\,\Phi\Big(\frac{\gamma}{\sqrt{V}} - \lambda\frac{Z}{\sqrt{V}}\Big) \end{aligned}$$

where Z and V are independent random variables and Z has a standard normal distribution and V is a chi-squared distribution with $n-1$ degrees of freedom divided by $n-1$ and

$$\gamma = \lambda \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}$$

Note that this expectation depends on the parameters θ and σ only through γ. To compute it we first condition on $V = v$.

$$\begin{aligned} E\,\Phi\Big(\frac{\gamma}{\sqrt{V}} - \lambda\frac{Z}{\sqrt{V}}\Big) &= E\,E\,\Phi\Big(\frac{\gamma}{\sqrt{V}} - \lambda\frac{Z}{\sqrt{V}} \mid V\Big) \\ &= E\,E\,\Phi\Big(b - aZ \mid V = v\Big) \end{aligned} \tag{11}$$

where

$$a = \lambda/\sqrt{v} \quad \text{and} \quad b = \gamma/\sqrt{v}$$

Let Z_1 and Z_2 be independent standard normal random variables. Then

$$\begin{aligned} E\,\Phi\Big(b - aZ \mid V = v\Big) &= E\,\Phi(b - aZ) \\ &= P(aZ_1 + Z_2 \le b) \\ &= \Phi\Big(\frac{b}{\sqrt{a^2+1}}\Big) \\ &= \Phi\Big(\frac{\lambda}{\sqrt{v+\lambda^2}}\frac{\theta_0-\theta}{\sigma/\sqrt{n}}\Big) \end{aligned}$$

Substituting the previous equation into equation 11 we see that we have proved the following lemma.

Lemma 2. *Let $\bar{X} = \sum_{i=1}^n X_i/n$, $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2/(n-1)$ and Φ be the distribution function of the standard normal distribution. Then*

$$E_{\theta,\sigma}\Phi\Big(\lambda\sqrt{n}\frac{\theta_0 - \bar{X}}{S}\Big) = E\,\Phi\Big(\frac{\lambda\sqrt{n}}{\sqrt{V+\lambda^2}}\frac{\theta_0-\theta}{\sigma}\Big) \tag{12}$$

where V is a chi-squared random variable with $n-1$ degrees of freedom divided by $n-1$.

The next step is to use the results of the lemma to find an approximate unbiased estimator of the fuzzy membership function given in equation 10. A simple Taylor series expansion about $E(V)=1$ for the expression in the right hand side of equation 12 gives the following.

$$E\,\Phi\Big(\frac{\lambda\sqrt{n}}{\sqrt{V+\lambda^2}}\frac{\theta_0-\theta}{\sigma}\Big) \doteq \Phi\Big(\frac{\lambda\sqrt{n}}{\sqrt{1+\lambda^2}}\frac{\theta_0-\theta}{\sigma}\Big) \tag{13}$$

If as in the previous section we let $a=\sqrt{n}\lambda/\sqrt{1+n^2}$ then the previous equation and the lemma yield

$$E_{\theta,\sigma}\Phi\Big(\frac{a}{\sqrt{1-a^2/n}}\frac{\theta_0-\bar{X}}{S}\Big) \doteq \Phi\Big(a\frac{\theta_0-\theta}{\sigma}\Big) \tag{14}$$

Simulation studies show that this approximation works quite well. We recall in passing that the best unbiased estimator of the equation 12 is well known when $a=1$. (See page 87 of Lehmann (1983).) In this case we compared our approximately unbiased estimator with the best unbiased estimator in a small simulation study with $n=5$ and observed that the two behave quite similarly.

In a particular problem to find an appropriate fuzzy membership function of the type in equation 10 we first select a value for θ_0. Next we select $0<\beta<0.5$, $\theta_1>\theta_0$ and $\sigma_1>0$ and solve equation 6 to get the value of a given in equation 7. This reflects our assessment of the point (θ_1,σ_1) belonging to the set of good parameter values.

If we are interested in unbiased estimators then we could consider fuzzy membership functions of the type in equation 9. For a fixed τ the function $\Phi((\theta_0-\theta)/\sqrt{\sigma^2/n+\tau^2})$ is easy to plot. The plot for the case when $\tau=5$,

$\theta_0 = 0$, $\theta \in (-5, 5)$ and $\sigma/\sqrt{n} \in (0, 25)$ is given in figure 1. Over this range of values for θ and $\sigma/\sqrt{n}$ as we increase the value of τ our contour lines will become nearly parallel and the plot almost flat. While if we let τ approach zero the lines will become rays emanating from zero and the surface will become much steeper.

One important way that this family of fuzzy membership functions is different from the family in equation 10 is that their values are no longer a function of standard units, that is a function of $(\theta_0 - \theta)/\sigma$. For some this could be a strong objection since the usual t-test depends on the parameter values only through their standardized value. But from the fuzzy point of view there seems to be no reason to impose this requirement.

3.5 The example again

To demonstrate our approach we return to the example in section 3.1. We must select the fuzzy membership function representing the set of good parameter values. In this case they are the parameter points (θ, σ) where the cholesterol of the population is of little concern. We begin letting $\theta_0 = 200$ which is a weak dividing line between the points of no concern and the rest of the parameter space. Next we select $\theta_1 = 215$ and decide we want our fuzzy membership function to have the value $\beta = 0.05$ at the point $(215, \sigma_1)$ for some choice of σ_1. The rationale behind choosing σ_1 is different than that for choosing $\theta_1 = 215$. This later choice is based on medical knowledge about the effects of cholesterol and does not depend on the true but unknown mean for this particular population. On the other hand our choice for σ_1 should be a reasonable guess for the standard deviation for the population at hand. For this example we will consider two possible choices for σ_1, 30 and 50. We can then use equation 10 and our choices for θ_1 and σ_1 to find the value of a to use in estimator 14. Similarly we use equation 9 to find the value of τ^2 for the estimator in equation 8.

The data collected by VA was a random sample of size 4921 with a sample mean of 210.9 and a standard deviation of 43.4. (For more information on the data see Rubins et al. (2003). For the two sets of two fuzzy membership functions we calculated the approximated unbiased estimators for the first set and the unbiased estimators for the second set. To help see the influence of sample size on our estimators we did this twice. Once for the true sample size of 4921 and a second time with sample size 200. The results are given in table 1.

We see from the table that both estimators are quite robust against sample size. The unbiased estimator of equation 8 is also robust against the choice of σ_1 while the approximate unbiased estimator of equation 14 is much more sensitive and should only be used when a good guess for the population standard deviation is available.

The usual p-value based on the t-test for $\theta \leq 200$ against the alternative $\theta > 200$ for our data is highly significant because of the large sample size. It is of little use here because what it is estimating is really of no interest. Why not then just estimate θ? The problem with this is that one would wishes to estimate the degree of membership of the unknown pair of parameter points (θ, σ) the set of good parameter points where the population's cholesterol is of little concern. This is not given by a point estimate of the population mean. Our approach requires one to choose a fuzzy membership function

n	σ_1	a	Est	τ	Est
4921	30	3.29	0.204	9.11	0.116
4921	50	5.50	0.084	9.09	0.115
200	30	3.29	0.198	8.87	0.101
200	50	5.50	0.068	8.41	0.097

Table 1: Values of the two fuzzy set estimators for the VA data for $\theta_0 = 200$, $\theta_1 = 215$, $\beta_1 = 0.05$, two choices of σ_1 and two choices of the sample size.

which models our levels of concern over the entire parameter space. Although not simple as the usual p-value it can be more informative.

4 Binomial problems

4.1 A family of fuzzy membership functions

We begin by recalling some facts about one sided binomal testing problems. Let X be binomial(n, θ) where n is known and $\theta \in [0, 1]$ is unknown and consider the testing problem

$$H : \theta \geq \theta_0 \quad \text{against} \quad K : \theta < \theta_0 \tag{15}$$

Let $P(X)$ denote the p-value coming from the UMP family of tests. If θ_0 is true and n is large then the distribution of $P(X)$ is approximately uniform on the unit interval and $E_{\theta_0} P(X)$ is approximately 0.5. $P(X)$ is essentially a smoother version of $1 - \phi(X, 0.5, \theta_0)$ and its expectation looks very much like $1 - E_\theta \phi(X, 0.5, \theta_0)$ for large n.

For ease of exposition we assume that θ is the proportion of patients which will respond to a new treatment. Let A denote the fuzzy set of useful treatments. For any value of θ the clinician needs to assess its degree of membership in this set. This value measures the overall desirability of the new treatment based on the current and perhaps somewhat limited information. This assessment depends on many factors such as its cost, ease of application, severity of side effects and so forth.

The first step in selecting a fuzzy membership function is choosing a value for θ_0, the "soft break" point between the useful values of θ and the rest of the parameter space. It will play a similar role to θ_0 in equation 15. In the case where we are considering a new treatment and there is a well accepted standard treatment we could take θ_0 to be the probability of a response under the standard treatment. However this need not be the case in general. If the new treatment could have less serious side effects, be easier to apply or be significantly cheaper then we could select a value for θ_0 which is less than the probability of response under the standard treatment.

For a positive integer $m < n$ let ϕ_m denote the UMP level 0.5 test of equation 15 based on Y_m a binomial(m, θ) random variable. Let $\lambda_m(\theta) = 1 - E_\theta \phi_m(Y_m)$. Then λ is a strictly increasing function on the unit interval whose range is also the unit interval and it takes on the value $1/2$ at $\theta = 1/2$. So each such function is a possible fuzzy membership function along with any finite convex combination of such functions. This is a reasonably rich family of functions which are easy to graph. In many problems it should not to difficult to select a sensible membership function.

After a sensible fuzzy membership function has been selected then one needs to find an estimator for it. It is well known (Lehman(1983)) that a function of θ has an unbiased estimator if and only if it is polynomial in θ of degree less than or equal to n. Clearly the family described just above have unbiased estimators. Finding the unbiased estimator of the selected fuzzy membership function is easy if we remember that the unbiased estimator of

$$\binom{m}{k}\theta^k(1-\theta)^{(m-k)}$$

is

$$\delta_{m,k}(x) = \begin{cases} 0 & \text{for } x < k \text{ or } x > n-(m-k), \\ \binom{m}{k}\binom{n-m}{x-k}/\binom{n}{x} & \text{for } k \le x \le n-(m-k). \end{cases}$$

4.2 An example

There has been some very recent interest in using Botox to relieve pain. See for example Singh et al. (2008) In a clinical trial 22 patients with chronic, refractory shoulder pain were injected with a mixture of Botox and lidocaine. After a month the patients were checked to see how many of them had experienced a meaningful reduction in their pain.

In the such clinical trials it is known that as many as 25% of the patients can experience a placebo effect. For this reason and the fact that little is known about the efficacy of Botox as a pain reliever we decide to use a soft break point of $\theta_0 = 0.35$. To choose an appropriate fuzzy membership function we considered convex mixtures of the UMP level 0.5 tests based on the sample sizes of 2, 7, 12, 17 and 21. In figure 2 the lines are the five fuzzy membership functions based on these tests. We see that all the fuzzy membership functions are approximately linear in the neighborhood of $\theta_0 = 0.35$. Hence, in this example, selecting a fuzzy membership function can come down to specifying its slope at $\theta_0 = 0.35$ and to a much lesser extent its behavior further away from this point. The question that needs to be addressed is how important are small differences in the neighborhood of $\theta_0 = 0.35$. The more important such differences are the steeper the fuzzy membership function should be around this point. For this problem the derivative of $1 - E_\theta \phi_m(Y_m)$ evaluated at $\theta = 0.35$ increases from 1.20 to 3.82 as m goes from 2 to 21. The curve represented by the small circles is the fuzzy membership function which is the convex mixture of these two with weights 0.7 on the test based on $m = 2$ and 0.3 on the test based on $m = 21$. Its slope at 0.35 is$.7 \times 1.20 + .3 \times 3.82 = 1.99$. The plot of the x's gives the values of its best unbiased estimator for a sample of size 22. In the actual trial 10 patients noted a reduction in their pain. The estimate of this fuzzy membership function for this outcome is 0.79 indicating strong evidence that the treatment belongs to the fuzzy set of useful treatments.

In figure 3 the two lines plot the expected value of the usual p-value and the fuzzy membership function described in the proceeding paragraph. For a sample of size $n = 22$ the circles plot the values of p-value and the x's plot the values of the unbiased estimator of our fuzzy membership function. The two curves are very similar. Remember however our fuzzy membership function was selected to represent the realities of a specific problem and does not depend on the sample size. If the sample size was increased however the curve of expected value of the p-value would change, getting steeper and steeper in the neighborhood of $\theta_0 = 0.35$. The p-value is designed to make as sharp of distinction as possible between values on the either side of θ_0. It is not clear to us why this is a good idea.

5 Concluding remarks

Our theory can be extended to other testing situations. For example the common two sample binomial and normal testing problems are easily handled.

Here we have focused on finding unbiased or approximately unbiased estimators of fuzzy membership functions as an alternative to computing p-values. For a Bayesian once the fuzzy membership function to be estimated has been selected and a prior chosen finding its Bayes estimator, in principle, is straightforward. The Bayesian approach always seems more natural in estimation than in testing. Our approach should work well and eliminate some of the problems associated with testing problems. Point null hypotheses have always been somewhat problematical for Bayesians. For example, Rousseau (2006) discusses a Bayesian approach where a point null is replaced by a small approximating interval hypothesis.

Some authors have considered testing hypotheses where the null and alternative are both described by fuzzy membership functions. These functions usually are piecewise linear. In such a setup they develop an analog of the Neyman-Pearson theory which is quite different from the estimation approach we have presented.

We have argued here that the usual Neyman-Pearson theory of hypotheses testing with the sharp division between the null and the alternative and accept-reject rules is not very useful in practice for many scientific questions. Moreover the usual p-value or level of significance does not really fix the problem. Our approach requires a careful assessment of the degree of membership for parameter points belonging to the special set of designated values. In selecting the appropriate fuzzy membership function more attention must be paid than when one is selecting the dividing point between the null and alternative hypotheses in standard methods. We belive that the payoff for the extra work is more useful inferences. We emphasize that there is nothing Bayesian in this. We are not assessing which are the likely or unlikely parameter values.

In their discussion of the notion of a level of significance Kempthorne and Folks (1971) emphasize that it is the ordering of the data values in strength of evidence against the null which is crucial. Once this is decided the rest follows easily. Note however in many problems the sensible order is usually obvious and hence there is only one sensible level of significance for a given data point once θ_0, the dividing point between the hypotheses, is selected. This suggests that the usual theory of p-values is to crude and does not allow for a more nuanced measure of evidence. Some might argue that this simplicity is in fact a strength of p-values. We disagree and believe that our approach allows for a more realistic measure of strength of evidence. We believe that if one has seriously contemplated the implications of various parameter values being true when selecting the fuzzy membership function to be estimated then the interpretation of the actual estimated value is easy and more informative.

References

Arnold, B. (1996). An approach to fuzzy hypotheses testing. *Metrika*, 44:119–126.

Blyth, C. and Staudte, R. (1995). Estimating statistical hypotheses. *Statitics and Probability Letters*, 23:45–52.

Casals, M., Gil, M., and Gil, P. (1986). On the use of Zadeh's probabilistic definition for testing statistical hypotheses from fuzzy information. *Fuzzy Sets and Systems*, 20:175–190.

Dollinger, M., Kulinskaya, E., and Staudte, R. G. (1996). Fuzzy hypothesis tests and confidence intervals. In Dowe, D., Korb, K., and Oliver, J., editors, *Information, Statistics and Induction in Science*, pages 119–128. World Scientific, Singapore.

Filzmoser, P. and Viertl, R. (2004). Testing hypotheses with fuzzy data: the fuzzy p-value. *Metrika*, 59:21–29.

Geyer, C. and Meeden, G. (2005). Fuzzy confidence intervals and P-values (with discussion). *Statistical Science*, 20:358–387.

Kempthorne, O. and Folks, L. (1971). *Probability, Statistics and Data Analysis*. Iowa state university press, Ames Iowa.

Klir, G. and St. Clair, U. (1997). *Fuzzy Set Theory: Foundations and Applications*. Prentice Hall PTR, Upper Saddle River, NJ.

Lehmann, E. (1983). *Theory of point estimation*. John Wiley & Sons, New York.

Lehmann, E. (1986). *Testing Statistical Hypotheses*. John Wiley & Sons, New York, second edition.

Manton, K., Woodbury, M., and Tolley, H. (1994). *Statistical Applications Using Fuzzy Sets*. John Wiley & Sons, New York.

Rousseau, J. (2006). Approximating interval hypothesis: p-values and Bayes factors. In *Proceedings of the Valencia/ ISBA 8th World Meeting on Bayesian Statistics*.

Rubins, H., Nelson, D., Noorbaloochi, S., and Nugent, S. (2003). Effectiveness of lipid lowering medications in outpatients with coronary heart disease in the department of veterans affairs system. *American Journal of Cardiology*, 92:1177–1192.

Singh, J., Mahowald, M., and S., N. (2008). Intra-articular botulinum toxin type a (ia-bont/a) significantly decreases shoulder pain in patients with refractory shoulder pain due to osteoarthritis: A randomized double-blind placebo controlled trial. *Arthritis and Rheumatism*, 56 (12):4233–4233.

Singpurwalla, N. and Booker, J. (2004). Membership functions and probability measures of fuzzy sets (with discussion). *Journal of the American Statistical Association*, 99:867–889.

Taheri, S. (2003). Trends in fuzzy sets. *Austrian Journal of Statistics*, 32:239–257.

Taheri, S. and Behboodian, J. (1999). Neyman-Pearson lemma for fuzzy hypoteses testing. *Metrika*, 49:3–17.

Wellek, S. (2003). *Testing Statistical Hypotheses of Equivalence*. Chapman & Hall/CRC, London.

Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8:338–359.

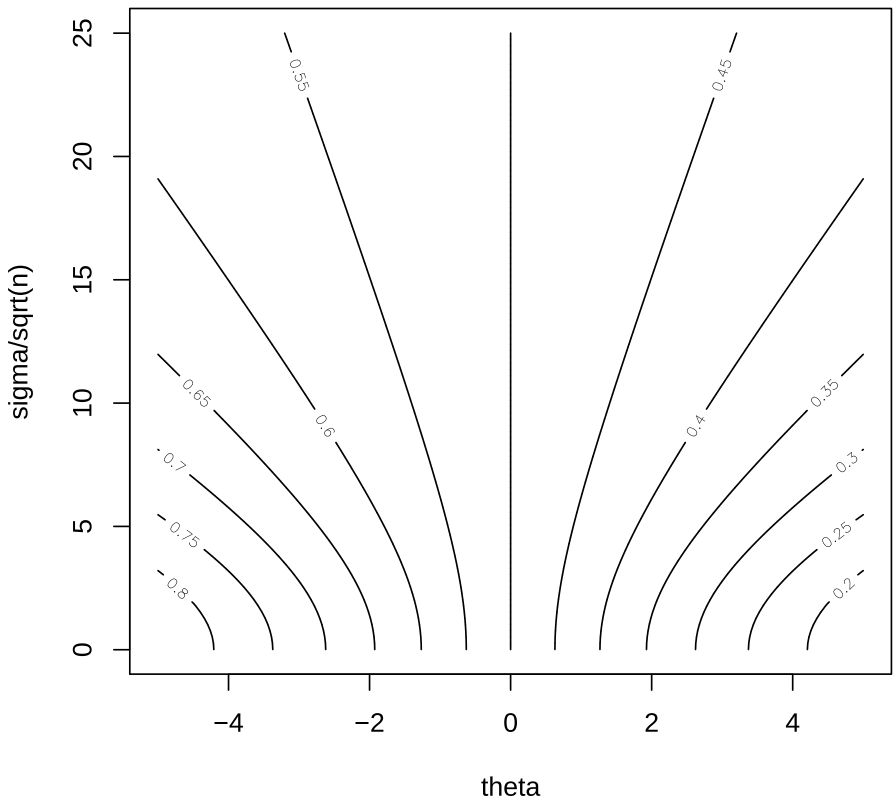

Figure 1: Contour plot of the fuzzy membership defined by $\tau = 5$ in equation 9

.

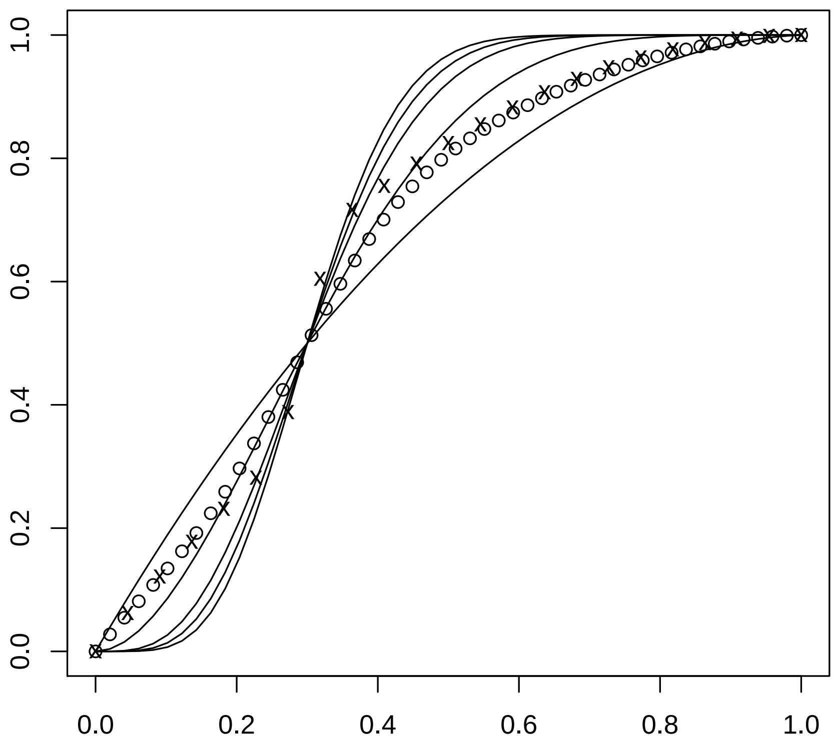

Figure 2: For the binomial example the lines are 5 possible fuzzy membership functions. The circles are a convex combination of 2 of them and the x's the estimates of this function for a sample of size $n = 22$.

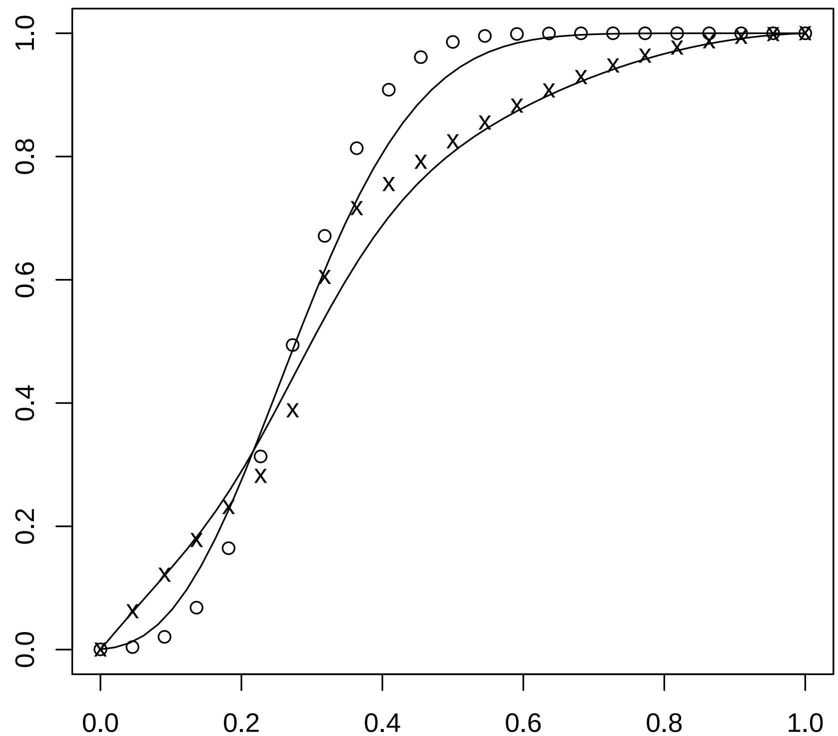

Figure 3: Plots of the expected value of the p-value and the fuzzy membership function in the binomial example. The circles are the values of the p-value and the x's are the estimates of the fuzzy membership function for a sample of size $n = 22$.