Comparing Two Tests used for Diagnostic or Screening Purposes

by

Seymour Geisser
University of Minnesota

Technical Report No. 613 (revised)
September 1996

Abstract

We present a Bayesian approach to the comparison of two binary diagnostic tests in a decision framework. It is shown, under a sensible loss function, that both a larger sensitivity and specificity are sufficient for one test to be superior to another. However, even if both the predictive value positive and negative of one test are larger than another test, it need not be superior, with respect to that loss function. An example of this approach is given that compares two Elisa screening tests for antibodies to the AIDS virus. We also show how a binary test that depends on a random variable may be optimally dichotomized with respect to the loss function.

Keywords: Bayes, decision framework, diagnostic test, dichotomization, predictive value negative, predictive value positive, screening test

Comparing Two Tests used for Diagnostic or Screening Purposes

Seymour Geisser, University of Minnesota

1 Introduction

Diagnostic tests for drug use, for antibodies to various viri and for diagnosing a variety of conditions, are increasingly being used. Gastwirth (1987) studied the accuracy rates of binary tests. Johnson and Geisser (1992) studied the optimal administration of dual tests from a Bayesian perspective. Here interest is focussed on comparing two imperfect tests, whether used for diagnosis or screening, for the same characteristic using a Bayesian approach. In order to accomplish this we first will assume that each subject to whom test i is administered will be designated as T_i or $\overline{T}_i$ indicating the presence of C or absence $\overline{C}$ of a characteristic or condition. In some tests, a putative diagnostic test is used as a screen for a condition in which the prevalence is low. When the screening test indicates the condition, a much more conclusive test, considerably more expensive, will then be applied to confirm or deny the existence of the conditions. In particular, such tests are most useful as indicators of the condition prior to the manifestation of overt symptoms or any easily recognizable features of the condition. We then define the following probabilities:

- $\Pr(C) = \pi$, the probability that a randomly drawn individual from a well defined population exhibits characteristic C is called the prevalence;
- $\Pr(T_i|C) = \eta_i$, the probability that test i correctly asserts the presence of C, is called the sensitivity;
- $\Pr(\overline{T}_i|\overline{C}) = \theta_i$, the probability that the test i correctly asserts the absence $\overline{C}$ is called the specificity;
- $\Pr(C|T_i) = \pi\eta_i/\{\pi\eta_i + (1-\pi)(1-\theta_i)\} = \psi_i$, the probability that the characteristic is present given test i indicates its presence is called the predictive value positive (pvp).
- $\Pr(\overline{C}|\overline{T}_i) = (1-\pi)\theta_i/\{\pi(1-\eta_i)+(1-\pi)\theta_i\} = \overline{\psi}_i$, the probability that the characteristic is absent given test i indicates its absence is called the predictive value negative (pvn).

Given the values π, η_i, and θ_i, the last two probabilities are obtained as a simple application of Bayes theorem. Generally speaking, we would say that test i dominates test t if

$$\psi_i > \psi_t \text{ and } \overline{\psi}_i > \overline{\psi}_t, \tag{1.1}$$

which is clearly the case when $\eta_i > \eta_t$ and $\theta_i > \theta_t$. Note that this is a sufficient but not a necessary condition for (1.1) to hold. As a counterexample, note that for $\eta_i = .1$, $\eta_t = .4$, $\theta_i = .9$, $\theta_t = .5$, $\psi_i = \frac{1}{2} > \frac{4}{9} = \psi_t$ and $\overline{\psi}_i = \frac{1}{2} > \frac{5}{11} = \overline{\psi}_t$, when $\pi = \frac{1}{2}$.

We may, however, introduce losses for particular actions. For example, if an individual is T_i, then assume action j will be taken and if $\overline{T}_i$, action k. Hence, we compute the loss in

taking action j for test i having observed T_i. Let l_{jC} be the loss in taking action j when C holds and $l_{j\overline{C}}$ when $\overline{C}$ holds, then

$$\begin{aligned} L_j(T_i) &= l_{jC}\Pr[C|T_i] + l_{j\overline{C}}\Pr(\overline{C}|T_i) \\ &= \psi_i[l_{jC} - l_{j\overline{C}}] + l_{j\overline{C}}. \end{aligned} \tag{1.2}$$

Similarly, when $\overline{T}_i$ is observed and l_{kC} and $l_{k\overline{C}}$ are respectively the losses for taking action k when C and $\overline{C}$ hold, the expected loss for a negative result is

$$\begin{aligned} L_k(\overline{T}_i) &= l_{kC}\Pr(C|\overline{T}_i) + l_{k\overline{C}}\Pr(\overline{C}|\overline{T}_i) \\ &= \overline{\psi}_i[l_{k\overline{C}} - l_{kC}] + l_{kC}. \end{aligned} \tag{1.3}$$

Note in comparing two tests i and t if (1.1) holds, then $L_j(T_i) < L_j(T_t)$ and $L_k(\overline{T}_i) < L_k(\overline{T}_t)$ whenever the set of reasonable assumptions $l_{jC} < l_{j\overline{C}}$ and $l_{k\overline{C}} < l_{kC}$ holds, i.e. the correct decisions are "less costly" than incorrect ones. Hence a sufficient condition for both conditional losses of test i to be less than those of test t given that the results of both tests are the same is (1.1).

Now to compute the overall expected loss for test i, we calculate

$$\begin{aligned} L_i &= L_j(T_i)\Pr(T_i) + L_k(\overline{T}_i)\Pr(\overline{T}_i) \\ &= [L_j(T_i) - L_k(\overline{T}_i)]\Pr(T_i) + L_k(\overline{T}_i) \end{aligned} \tag{1.4}$$

and the superior test is the one associated with the the smaller of L_i or L_t noting we have made the reasonable assumption that the same action is taken for every individual and depends only whether an individual is T or $\overline{T}$ irrespective of the test.

If we assume zero loss for correct decisions so that $l_{jC} = l_{k\overline{C}} = 0$ and unit loss for one incorrect decision and a multiple $a > 0$ for the other, then we can rewrite (1.4) as

$$\begin{aligned} L_i &= (1 - \psi_i)\Pr(T_i) + a(1 - \overline{\psi}_i)\Pr(\overline{T}_i) \\ &= (1 - \pi)(1 - \theta_i) + a\pi(1 - \eta_i). \end{aligned} \tag{1.5}$$

Therefore, for $a > 0$, a sufficient condition for

$$L_i < L_t \tag{1.6}$$

is still as before

$$\eta_i > \eta_t \text{ and } \theta_i > \theta_t. \tag{1.7}$$

However, (1.1) is not sufficient for (1.6), see either the previous counterexample, or note that $\Pr(T)$ depends on the test being used and the prevalence. Another way of looking at this is that when the tests give the same outcome, test i is superior but this is not true when the tests yield different outcomes.

More generally, when the parameters π, η_i, and θ_i are unknown but can be estimated from data via a Bayesian approach, the probabilities ψ_i and $\overline{\psi}_i$ are random variables conditioned

on the data d. One assumes a joint prior $p(\pi, \eta_1, \eta_2, \theta_1, \theta_2)$ and combines this with a likelihood from a previous experiment $L(\pi_1, \eta_1, \eta_2, \theta_1, \theta_2|d)$ to obtain

$$p(\pi, \eta_1, \eta_2, \theta_1, \theta_2|d). \tag{1.8}$$

From the above, one then calculates

$$\Pr(C|T_i, d);\ \Pr(C|\overline{T}_i, d);\ P(T_i|d) \tag{1.9}$$

and substitutes them in place of

$$\Pr(C|T_i),\ \Pr(C|\overline{T}_i),\ P(T_i) \tag{1.10}$$

in (1.2) - (1.6) as indicated in Geisser and Johnson (1992).

2 Use of Previous Studies

The posterior distribution given as (1.5) depends on how the likelihood was generated and how the prior distribution was assigned to the set of parameters.

Regarding the likelihood, there are several ways previous studies on the set of parameters could have been made, c.f. Gastwirth (1987), Geisser and Johnson (1992). First, we suppose a separate study on the prevalence, i.e. an independent random sample from the target population of sample size v was obtained and the number t_c having C is observed. Hence the likelihood for π is given as

$$L(\pi) \propto \pi^{t_c}(1-\pi)^{v-t_c}. \tag{2.1}$$

If separate studies for the parameters of tests 1 and 2, (η_1, θ_1) and (η_2, θ_2), using independent random samples of size $n_1, n_2, \overline{n}_1, \overline{n}_2, n_i$ drawn from individuals with C and $\overline{n}_i$ with $\overline{C}$ were performed then the joint likelihood would be, for r_i testing T_i and $\overline{r}_i$ testing $\overline{T}_i$,

$$\mathcal{L}_1 = \prod_{i=1}^{2} L(\eta_i, \theta_i) = \prod_{i=1}^{2} \eta_i^{r_i}(1-\eta_i)^{n_i-r_i}\theta_i^{\overline{r}_i}(1-\theta_i)^{\overline{n}_i-\overline{r}_i}. \tag{2.2}$$

There are cases where the likelihood reflects only those assigned to t_c who overtly manifest the condition. Unless a so-called "gold standard" is used, π may be underestimated if there is a significant lag between having the condition and overt features being recognizable. In such cases two independent random samples of size s_i, $i = 1, 2$, from the population could be obtained. In these samples one would ascertain the number t_i that yield T_i. Since

$$\Pr(T_i) = \pi\eta_i + (1-\pi)(1-\theta_i)$$

the likelihood from this study is

$$\mathcal{L}_2 = L(\pi, \eta_1, \eta_2, \theta_1, \theta_2) \propto \prod_{i=1}^{2}[\pi\eta_i + (1-\pi)(1-\theta_i)]^{t_i}[\pi(1-\eta_i) + (1-\pi)\theta_i]^{s_i-t_i}. \tag{2.3}$$

Depending then on what previous studies have been done we could calculate the posterior distribution for $\theta = (\theta_1, \theta_2)$, $\eta = (\eta_1, \eta_2)$

$$p(\eta, \theta, \pi|d) \propto L(\pi)\mathcal{L}_1 p(\eta, \theta, \pi) \tag{2.4}$$

or

$$p(\eta, \theta, \pi|d) \propto \mathcal{L}_2\mathcal{L}_1 p(\eta, \theta, \pi). \tag{2.5}$$

Another possibility which requires joint testing, Geisser and Johnson (1992), can be applied. If those known to be C and $\overline{C}$ can be sampled from, then a joint study can be made where Table 1 reflects the various probabilities for both tests being applied.

Table 1

	C			$\overline{C}$	
	T_2	$\overline{T}_2$		T_2	$\overline{T}_2$
T_1	η_{11}	η_{10}	T_1	θ_{11}	θ_{10}
$\overline{T}_1$	η_{01}	η_{00}	$\overline{T}_1$	θ_{01}	θ_{00}

where, say, $\Pr[T_1, T_2|C] = \eta_{11}$, $\Pr[\overline{T}_1, T_2|\overline{C}] = \theta_{01}$, etc. We assume that n and $\overline{n}$ individuals are independently sampled at random from C and $\overline{C}$. Then

$$\mathcal{L}_3 = \prod_{i,j=0}^{1} \eta_{ij}^{r_{ij}} \theta_{ij}^{\overline{r}_{ij}} \tag{2.6}$$

where $\sum_{i,j} r_{ij} = n$ and $\sum_{i,j} \overline{r}_{ij} = \overline{n}$. Note $\eta_1 = \eta_{11} + \eta_{10}$, $\eta_2 = \eta_{11} + \eta_{01}$, $\theta_1 = \theta_{00} + \theta_{01}$, $\theta_2 = \theta_{00} + \theta_{10}$.

Another possibility for joint testing would be to obtain a random sample of size t and apply both tests to each individual and sort out the individuals according to categories T_1T_2, $T_1\overline{T}_2$, etc. so that, e.g.,

$$\Pr(T_1, T_2) = \pi\eta_{11} + (1 - \pi)\theta_{11} \tag{2.7}$$

and

$$\Pr(\overline{T}_1, T_2) = \{\pi\eta_{01} + (1 - \pi)\theta_{01}\}, \tag{2.8}$$

etc. Hence the likelihood here is

$$\mathcal{L}_4 = \prod_{i,j=0}^{1} \{\pi\eta_{ij} + (1 - \pi)\theta_{ij}\}^{t_{ij}} \tag{2.9}$$

where t_{ij} is the number of individuals in cell i, j.

Various combinations of these likelihoods, depending on what data are available, or what studies might be contemplated to be undertaken, can be used with the prior to obtain posteriors.

3 The Comparison

Recall the initial comparison is the set $(\psi_1, \overline{\psi}_1)$ with $(\psi_2, \overline{\psi}_2)$. When these are random variables given d then our focus could be any of the following for $i = 1, 2$:

$$E(\psi_i|T_i, d) = \Pr[C|T_i, d] = \frac{E(\pi\eta_i|d)}{E(\pi\eta_i|d) + E[(1-\pi)(1-\theta_i)|d]} \tag{3.1}$$
$$E(\overline{\psi}_i|\overline{T}_i, d) = \Pr[\overline{C}|\overline{T}_i, d] = \frac{E[\theta_i(1-\pi)|d]}{E[\theta_i(1-\pi)|d] + E[\pi(1-\eta_i)|d]}$$

so, based on this particular characteristic, we could consider test 1 weakly dominating test 2 if

$$\Pr[C|T_1, d] > \Pr[C|T_2, d] \tag{3.2}$$

and

$$P[\overline{C}|\overline{T}_1, d] > \Pr[\overline{C}|\overline{T}_2, d]. \tag{3.3}$$

Note that (3.2) and (3.3) will hold if the stronger conditions

$$E(\eta_1|d) > E(\eta_2|d) \tag{3.4}$$
$$E(\theta_1|d) > E(\theta_2|d)$$

hold. The rhs of (3.1) also represents the predictive probability that an individual having tested $T(\overline{T})$ will have $C(\overline{C})$. When choosing a diagnostic test, we could prefer test 1 over test 2 if marginally

$$\Pr[\psi_1 \geq \psi_2|d] > p > \frac{1}{2} \tag{3.5}$$
$$\Pr[\overline{\psi}_1 \geq \overline{\psi}_2|d] > p > \frac{1}{2}$$

or jointly

$$\Pr[\psi_1 \geq \psi_2, \overline{\psi}_1 \geq \overline{\psi}_2|d] \geq q > \frac{1}{4}, \tag{3.6}$$

where q is the maximum of the 4 possible probabilities. Further, if appropriate loss functions are introduced, then the superiority of one test over another will depend on an analysis in (1.4). These calculations will depend on the prior and which studies were actually used.

For the sampling scheme which resorts to the direct information on π, namely (2.1) combined with either $\mathcal{L}_1$ or $\mathcal{L}_3$ and independent Dirichlet priors for $\{\eta_{ij}\}$ $\{\theta_{ij}\}$ and π we can calculate the expectations in (3.1), as

$$E(\psi_i) = \Pr[C|T_i, d] = \frac{E(\pi|d)E(\eta_i|d)}{E(\pi|d)E(\eta_i|d) + E(1-\pi|d)E(1-\theta_i|d)} \tag{3.7}$$
$$E(\overline{\psi}_i) = \Pr[\overline{C}|\overline{T}_i, d] = \frac{E(1-\pi|d)E(\theta_i|d)}{E(1-\pi|d)E(\theta_i|d) + E(\pi|d)E(1-\eta_i|d)}.$$

To calculate the expected losses of (1.2) - (1.4) we substitute for ψ_i and $\overline{\psi}_i$, $\Pr(C|T_i, d)$ and $\Pr(\overline{C}|\overline{T}_i, d)$ and $\Pr(T_i|d)$ for $\Pr(T_i)$ and it is sufficient for

$$L_1(d) < L_2(d) \tag{3.8}$$

if (3.4) holds.

4 An Example

In Geisser and Johnson (1992), two Elisa tests for the AIDS virus (actually antibodies to the virus) were considered. One test was Dupont (1) and the other was Abbot (2). The relevant statistics are given there and encapsulated here:

$$\begin{aligned} E(\eta_1|C,d) &= .984, \quad & E(\eta_2|C,d) &= .994 \\ E(\theta_1|\overline{C},d) &= .958, \quad & E(\theta_2|\overline{C},d) &= .924 \end{aligned} \tag{4.1}$$

$$E(\pi|d) = .000159$$

$$\begin{aligned} E(\psi_1|d) &= .0035 \quad & E(\overline{\psi}_1|d) &= .9999976 \\ E(\psi_2|d) &= .0019 \quad & E(\overline{\psi}_2|d) &= .9999992, \end{aligned} \tag{4.2}$$

based on the beta priors used there for $\eta_1, \eta_2, \theta_1, \theta_2$ and π. Calculations for the overall losses for the two tests are

$$\begin{aligned} L_1 &= .04199332 + .000002544a \\ L_2 &= .075987916 + .000000954a \\ L_2 - L_1 &= .033994596 - .00000159a. \end{aligned} \tag{4.3}$$

Therefore

$$L_2 > L_1 \text{ for } a < 21,380.25. \tag{4.4}$$

Hence for $a < 21,380.25$ Dupont is superior, otherwise Abbot is.

Now if an Elisa test (either Dupont's or Abbot's) yields a positive result, the individual would be subjected to a further confirmatory test, usually the Western Blot, because of the small expected pvp in both tests. Here both tests are basically used as screening devices. Hence a false positive requires the unit cost of a further confirmatory test. However, a false negative could be much more serious since an individual presumably would not be treated and could communicate the virus to others. Hence if the cost of a false negative were less than 21,380.25 units of the cost of the confirmatory test then the Dupont is superior, otherwise the Abbot is. Of course, all of this depends on this particular experiment, Burkhardt et al. (1987) and the use by Geisser and Johnson (1992) on particular priors and the Canadian data used, Nusbacher et al. (1986). Hence this should only be taken as an illustration of the methodology rather than as a conclusion concerning the two tests. Clearly, as purely

diagnostic tests for the presence of the AIDS virus, the expected pvp is miniscule for both tests, and their virtue lies in their screening capacity. A diagnostic for a condition will be useful if

$$\begin{aligned} E(\psi|d) &> \max(.5, E(\pi|d)) \\ E(\overline{\psi}|d) &> \max(.5, 1 - E(\pi|d)). \end{aligned} \tag{4.5}$$

For other criteria of diagnostic utility, see Viana and Farewell (1990, 1994). However, since these tests have an extremely high expected pvn, the usefulness is vested in an inexpensive screening for AIDS in the sense that a positive on a test is retested with a "confirmatory" test with a very high pvp such as the much more expensive Western Blot.

It is to be noted that the Canadian data represented a low risk group. On the other hand, it has been estimated that among gay males in the United States, $E(\pi|d) = .07$. Use of this yields $L_2 > L_1$ for $a < 45.17$ and $L_1 > L_2$ otherwise, which could considerably change our view of which test to use.

5 Dichotomization of a Diagnostic Test Variable

In many situations, a diagnostic or screening test depends on a random variable Y, say, that is dichotomized to result in a binary test. In these cases where η is, say, an increasing function of Y and θ decreasing in Y, we write, for distribution functions F and G,

$$\begin{aligned} \Pr(T|C) &= \eta = F(y) \\ \Pr(\overline{T}|\overline{C}) &= \theta = 1 - G(y). \end{aligned} \tag{5.1}$$

We then obtain the optimal value $\hat{y}$ that minimizes the loss function as given in (1.5). Hence

$$\begin{aligned} L(y) &= [(1-\pi)G(y) + a\pi(1 - F(y))] \\ \hat{y} &= \arg\min_y L(y). \end{aligned} \tag{5.2}$$

For example, if the distribution functions are both simple exponentials so that

$$\begin{aligned} F(y) &= 1 - e^{-\lambda y} \\ G(y) &= 1 - e^{-\mu y} \end{aligned}$$

then

$$L(y) = (1-\pi)(1 - e^{-\mu y}) + a\pi e^{-\lambda y}.$$

Minimization of $L(y)$ results in

$$\hat{y} = \max\left(\frac{1}{\mu - \lambda} \log \frac{(1-\pi)\mu}{\pi a \lambda}, 0\right). \tag{5.3}$$

This assumes that π, μ and η are known. When there is independent information on π whose likelihood can be combined with independent information regarding $F(y)$ and $G(y)$ we can proceed as follows. Assuming that F and G have known forms but unknown parameters,

we can use a Bayesian approach to calculate the expected loss and then minimize it with respect to y. In the previous exponential example, suppose there are independent random samples of size N from C and M from $\overline{C}$ on F and G respectively. If we assume independent "non-informative priors" $p(\mu, \lambda) = \frac{1}{\lambda\mu}$ on μ and λ, we obtain predictive distributions

$$\tilde{F}(y|d) = 1 - \left(1 + \frac{y}{N\overline{y}_1}\right)^{-N} \tag{5.4}$$
$$\tilde{G}(y|d) = 1 - \left(1 + \frac{y}{M\overline{y}_2}\right)^{-M},$$

where $\overline{y}_1$ and $\overline{y}_2$ are the sample averages from the random samples on F and G. Hence we minimize the expected loss

$$\overline{L}(y|d) = (1 - \overline{\pi})\left[1 - \left(1 + \frac{y}{N\overline{y}_1}\right)^{-N}\right] + a\overline{\pi}\left(1 + \frac{y}{M\overline{y}_2}\right)^{-M} \tag{5.5}$$

where $\overline{\pi} = E(\pi|d)$. Hence setting

$$\frac{d\overline{L}(y|d)}{dy} = \frac{(1-\overline{\pi})}{\overline{y}_1}\left(1 + \frac{y}{N\overline{y}_1}\right)^{-(N+1)} - \frac{a\overline{\pi}}{\overline{y}_2}\left(1 + \frac{y}{N\overline{y}_2}\right)^{-(M+1)} = 0,$$

numerical solutions for $\hat{y}$ can be found. For the simple case $N = M$, the exact algebraic solution is

$$\hat{y} = \max\left(\frac{N(B-1)\overline{y}_1\overline{y}_2}{\overline{y}_2 - \overline{y}_1 B}, 0\right) \tag{5.6}$$

where

$$B^{N+1} = \frac{(1-\overline{\pi})\overline{y}_2}{\overline{y}_1 a\overline{\pi}}.$$

Once the value $\hat{y}$ is obtained, then

$$E(\eta|d, \hat{y}) = 1 - \left(1 + \frac{\hat{y}}{N\overline{y}_1}\right)^{-N} = \overline{\eta}(\hat{y}) \tag{5.7}$$
$$E(\theta|d, \hat{y}) = \left(1 + \frac{\hat{y}}{M\overline{y}_2}\right)^{-M} = \overline{\theta}(\hat{y})$$

and

$$E(\psi|d) = \frac{\overline{\pi}\overline{\eta}(\hat{y})}{\overline{\pi}\overline{\eta}(\hat{y}) + (1-\overline{\pi})(1-\overline{\theta}(\hat{y}))} \tag{5.8}$$
$$E(\overline{\psi}|d) = \frac{(1-\overline{\pi})\,\overline{\theta}(\hat{y})}{(1-\overline{\pi})\,\overline{\theta}(\hat{y}) + \overline{\pi}(1-\overline{\eta}(\hat{y}))}.$$

In general, if F and G are known parametric distributions indexed by unknown $\alpha_1 \in A_1$, $\alpha_2 \in A_2$ and independently-distributed observations $y^{(N_k)} = (y_{k1}, \ldots y_{kN_k})$, $k = 1, 2$ are available on $F_1(y|\alpha_1)$ and $F_2(y|\alpha_2)$ where $F_1 = F$ and $F_2 = G$ then

$$\mathcal{L}(\alpha_k) = \prod_{i=1}^{N_k} f_k(y_{ki}|\alpha_k) \quad k = 1, 2. \tag{5.9}$$

Let the prior for α_1 and α_2 be $p(\alpha_1, \alpha_2) = p(\alpha_1)\, p(\alpha_2)$. Then

$$\tilde{F}_k\left(y|y^{(N_k)}\right) \propto \int_{A_k} \mathcal{L}(\alpha_k) F_k(y|\alpha_k)\, p(\alpha_k) d\alpha_k \tag{5.10}$$

for $k = 1, 2$. This results in

$$\overline{L}(y|d) = (1 - \overline{\pi})\tilde{F}_2\left(y|y^{(N_1)}\right) + a\overline{\pi}\left(1 - \tilde{F}_1\left(y|y^{(N_2)}\right)\right). \tag{5.11}$$

Optimization then yields

$$\hat{y} = \arg\min \overline{L}(y|d), \tag{5.12}$$

resulting in $\hat{y}$ as the optimal dichotomization of the test, which is inserted into (5.11). The choice of the superior of the two tests depends then on comparisons of $\overline{L}_i(\hat{y}_i|d)$, calculated on test i, and $\overline{L}_t(\hat{y}_t|d)$, calculated for test t.

Assuming Poisson distributions for F_1 and F_2 and samples of size N_1 and N_2 with priors α_1^{-1} and α_2^{-1}, we obtain negative binomial predictive distributions resulting in

$$\begin{aligned} \overline{L}(y|d) = (1 - \overline{\pi}) \sum_{t=0}^{y} \binom{N_1\overline{y}_1 + t - 1}{t} \left(\frac{1}{N_1 + 1}\right)^t \left(\frac{N_1}{N_1 + 1}\right)^{N_1\overline{y}_1} \\ + a\overline{\pi} \sum_{t=y+1}^{\infty} \binom{N_2\overline{y}_2 + t - 1}{t} \left(\frac{1}{N_2 + 1}\right)^t \left(\frac{N_2}{N_2 + 1}\right)^{N_2\overline{y}_2}. \end{aligned} \tag{5.13}$$

Optimization requires minimizing (5.13) with respect to y. Some simplification occurs when $N_1 = N_2$.

Similarly for F_1 and F_2 normally distributed and the usual non-informative prior $p(\mu_1, \mu_2, \sigma_1, \sigma_2) \propto \frac{1}{\sigma_1\sigma_2}$ we obtain

$$\frac{d\overline{L}(y|d)}{dy} = (1 - \overline{\pi})\tilde{f}_1 - a\overline{\pi}\tilde{f}_2 \tag{5.14}$$

where

$$\tilde{f}_i(y) = \frac{\Gamma(N_i/2)}{\Gamma\left(\frac{N_i - 1}{2}\right)} \left(\frac{s_i^2(N_i + 1)}{\pi(N_i - 1)N_i}\right)^{\frac{1}{2}} \left(1 + \frac{(y - \overline{y})^2 N_i}{(N_i^2 - 1)s_i^2}\right)^{-\frac{N_i}{2}}. \tag{5.15}$$

Setting

$$(1 - \overline{\pi})\tilde{f}_1(y) = a\pi\tilde{f}_2(y) \tag{5.16}$$

and solving for y numerically will yield a solution. Again, some simplification occurs when $N_1 = N_2$.

These 3 cases, exponential, Poisson and normal are, of course, easily extended to their conjugate priors when prior information can be put into that format. In particular cases where subjective priors are available on the parameters involved then they of course should be used.

If $\eta(\mathbf{x})$ and $\theta(\mathbf{x})$ are functions of a vector covariate $\mathbf{x} = (x_1, \ldots, x_p)'$ and distributed as known $F = F_1(y|\mathbf{x}, \beta_1, \sigma_1^2)$ and $G = F_2(y|\mathbf{x}, \beta_2, \sigma_2^2)$ such as normal linear regression, logistic regression or Poisson regression, the dichotomizer, for T or $\overline{T}$ would now be a variable $\hat{y}(\mathbf{x})$. Here the testees are no longer regarded as exchangeable and in order to reduce complexity one requires that either π be known or, if unknown, is assessed independently of θ and η, preferably by a presumed gold standard, e.g. Western Blot for AIDS. The Elisa test measures the level of certain antigens in the blood and a cutoff point would need to be set. The level of antigens in the blood may depend on several covariates so that the optimal dichotomizer for an individual may depend on the individual's covariates and yield $\hat{y}(\mathbf{x})$, Wittes (1987).

Further, when F_1 and F_2 have unknown parameters, then one would use the predictive distribution $\tilde{F}_1(y|\mathbf{x}, Y_1^{(N_1)}, X_1^{(N_1)})$ and $\tilde{F}_2(y|\mathbf{x}, Y_2^{(N_2)}, X_2^{(N_2)})$ where $Y_k^{(N_k)} = (y_{k1}, \ldots, y_{kN_k})$ and $X_k^{(N_k)} = (\mathbf{x}_{k1}, \ldots, \mathbf{x}_{kN_k})$ is the matrix of covariates and $\mathbf{x}_{kj}$ is the vector of covariates for the jth individual, for suitable priors on β_k and σ_k^2, $k = 1, 2$. Since the optimal $\hat{y}$ depends on $\mathbf{x}$, a comparison of the two tests is suitable only for a particular testee. If only one test is to be used for all testees in a mass screening program, then the distribution of the covariates needs to be taken into account for the test choice. Suppose then for a known distribution of covariates $H(\mathbf{x})$, we can compute the expected loss over the distribution of $\mathbf{x}$, then from (5.11),

$$\overline{L}(d) = E_{\mathbf{x}}\left[(1-\overline{\pi})\tilde{F}_2\left(\hat{y}(X)|Y_1^{(N_1)}, X_1^{(N_1)}, \mathbf{x}\right) + a\overline{\pi}\left(1 - \tilde{F}_1\left(\hat{y}(x)|Y_2^{(N_2)}, X_2^{(N_2)}, \mathbf{x}\right)\right)\right]. \tag{5.17}$$

One then would use the test with the smaller overall expected loss (5.17).

6 Acknowledgment

This work was supported in part by NIH Grant GM25271.

References

Burkhardt, U., Mertens, Th., and Eggers, H.J. (1987). Comparison of two commercially available anti-HIV ELISAs: Abbott HLTV III EIAc and DuPont HTLV III-ELISAs. *Journal of Medical Virology* **23** 217–224.

Gastwirth, J.L. (1987). The statistical precision of medical screening tests. *Statistical Science* **2** 213–238.

Geisser, S., and Johnson, W. (1992). Optimal administration of dual screening tests for detecting a characteristic, with special relevance to low prevalence diseases. *Biometrics* **48** 839–852.

Nusbacher, J. Chiavetta, J., Naiman, R., Buchner, B., Scalia, V., and Horst, R. (1986). Evaluation of a confidential method of excluding blood donors exposed to human immunodeficiency virus. *Transfusion* **26** 539–541.

Viana, M. and Farewell, V. (1990). A test for diagnostic utility. *Canadian Journal of Statistics* **18** (4) 289–295.

Viana, M. and Farewell, V. (1994). Assessing the diagnostic utility of a test. *Biometrics Journal* **36** 131–145.

Wittes, J. (1987). Comment. *Statistical Science* **2** (3) 228–230.