To appear in *Risk Analysis—An International Journal.*

# Interpreting Null Results from Measurements with Uncertain Correlations: An Info-Gap Approach

Yakov Ben-Haim[1]

**Abstract**

Null events—not detecting a pernicious agent—are the basis for declaring the agent is absent. Repeated nulls strengthen confidence in the declaration. However, correlations between observations are difficult to assess in many situations and introduce uncertainty in interpreting repeated nulls. We quantify uncertain correlations using an info-gap model, which is an unbounded family of nested sets of possible probabilities. An info-gap model is non-probabilistic and entails no assumption about a worst case. We then evaluate the robustness, to uncertain correlations, of estimates of the probability of a null event. This is then the basis for evaluating a non-probabilistic robustness-based confidence interval for the probability of a null.

**Keywords:** Uncertain correlations, null results, info-gaps.

## 1 Introduction

"No news is good news" applies to situations where one has searched diligently for a pernicious agent and not found it. In monitoring programs for verifying the eradication of an invasive species (Regan *et al.* 2006, Moffitt *et al.* 2008, Rout *et al.* 2009) a long series of null measurements is taken as strong evidence of absence. In monitoring the spread of an infectious disease such as SARS, mad-cow disease or influenza, the absence of cases indicates actual absence of the disease from the monitored population. In port-of-entry monitoring for detecting weapons or explosives (Moffitt *et al.* 2005), "no news" suggests that deterrent measures have been effective.

No news *is* good news provided that the null measurements are sufficiently numerous and independent of each other. This requires knowledge of the *degree of correlation* among the measurements. If a null measurement enhances the probability that subsequent measurements will also be null (regardless of the objective situation), then lots of "no news" is little or no "news" at all.

Correlations among purportedly independent measurements can arise by many mechanisms.

Searches are often performed by humans, or at least with an essential human element. Both positive and negative correlations among measurements can result from a "self-fulfilling prophecy" (SFP) as we now elucidate. The "prophecy" does not have to be explicit or conscious; it can be subliminal or implicit.

Positive correlations between measurements are induced by SFPs such as "Everybody has found one, so we must find one." Over-reporting of rare species when monitoring the decline of a population can occur unless strict observational protocols are used (such as bringing in physical specimens and not relying on visual sightings). For instance, Franklin (1999) uses a range of observational data from many different sources over the past 150 years—of varying accuracy and reliability—to evaluate change in bird assemblages in northern Australia. Some of these sources were trained biologists, though professional protocols changed over the sampling period. Some observers were casual or untrained observers who may exert less effort, and thus miss the rare events, or who are enthusiastic in the search for rare occurrences and may systematically over-report extreme observations. While historical observational data are an important and valuable source, it is difficult to verify that the assumption of independence is not violated.



[1]Yitzhak Moda'i Chair in Technology and Economics, Faculty of Mechanical Engineering, Technion—Israel Institute of Technology, Haifa, Israel, yakov@technion.ac.il

Positive correlations can also be induced by SFPs such as "Nobody has found one, so there probably aren't any." Boredom, over-confidence and inter-personal dynamics are perpetual challenges in monitoring rare events. Redundancy of human inspection and supervision is a key tool in clinical safety (Vincent 2006). However, the "[q]uantification of error is the most difficult aspect" of human reliability analysis (Lyons *et al.* 2004, p.231). Correlations among observations by different individuals are very hard to measure, and may be either positive or negative depending on inter-personal relations.

Negative correlations can also be induced by SFPs. Thoughts such as "Everybody has already found one, so we needn't look too hard", or "Nobody has found one, so we'd better find one" will both result in negative correlations among measurements.

We have discussed mechanisms relating to the searcher, but SFPs of various sorts can act, either explicitly or implicitly, on the potential target.

Positive correlation among measurements can result from motivational reasoning by the target such as "They haven't found us yet and we sure won't let them find us now" or "They find us wherever we are. What's the use?". These mechanisms can operate even without volition on the part of the target if the search process itself either enhances or diminishes the vulnerability of the target. Searches which tend to flush out the target will have positive correlation, while searches which tend to drive the target away will induce negative correlation.

Likewise, negative correlations can result from SFPs in the target population such as "They have begun to find us. We must hide better". An example is landowners who preemptively destroy habitat for an endangered species in order to avoid potential land-use regulations prescribed under the Endangered Species Act (Lueck and Michael, 2003). The converse motivation may also operate on a target, such as "They have never found us. We needn't hide so well".

The point of this discussion is that prevention of correlations among measurements requires careful design and control of the measurement protocol and conditions. This in turn depends on knowledge of both the target and the monitor. In situations such as emerging diseases, intelligent adversarial targets, or human observation under stress or repetition, this knowledge may be deficient. Indeed, the monitoring may be intended to detect new and unknown threats. When this is the case, the interpretation of a string of null results must account for uncertainty in correlation between the measurements.

When this uncertainty can be modeled probabilistically then standard statistical tools such as Bayesian inference or confidence intervals can be used. However, we consider situations in which probabilistic information is deficient. Consequently we employ a non-probabilistic concept of robustness to ignorance.

'Robustness' has many meanings. As we will use it, the concept of robustness derives from a prior concept of non-probabilistic uncertainty. Knight (1921) elaborated on the distinction between 'risk' based on known probability distributions and 'true uncertainty' for which probability distributions are not known. We are concerned with robustness against Knightian uncertainty.

Wald (1945) studied the problem of statistical hypothesis testing based on a random sample whose probability distribution is not known, but whose distribution is known to belong to a given class of distribution functions. Wald states that "in most of the applications not even the existence of ... an a priori probability distribution [on the class of distribution functions] ... can be postulated, and in those few cases where the existence of an a priori probability distribution ... may be assumed this distribution is usually unknown." (p.267). Wald introduced a loss function expressing the "relative importance of the error committed by accepting" one hypothesized subset of distributions when a specific (though unknown) distribution in fact is true. (p.266). He notes that "the determination of the [loss function] is not a statistical question and is considered here as given." (p.266). Wald developed a decision procedure which "minimizes the maximum ... of the risk function." (p.267).

Many engineering researchers, beginning in the 1960s, developed estimation and control algorithms for linear dynamic systems based on sets of inputs. Schweppe (1973) for instance develops inference and decision rules based on assuming that the uncertain phenomenon can be quantified in such a way as to be bounded by an ellipsoid, with no probability function involved.

This paper uses info-gap theory which builds on this tradition of non-probabilistic robustness concepts.

Our aim is to estimate the probability of absence based on a series of null results, and to rule out values less than the estimate at specified level of significance. Many statistical methods have been proposed from this task. However, as pointed out by Coolen (2006), severely uncertain data may not point clearly either to a single estimate or a single estimation algorithm. Coolen then proposes the method of non-parametric predictive inference to obtain bounds on the estimate.

Nonetheless it is sometimes necessary to obtain a point estimate of the probability of a null, and to exclude values less than the estimate. We use a binary hypothesis test for this purpose. However, we are uncertain that the measurements are statistically independent, and we observe only null events. Consequently we are unable to estimate both $p$ and the strength of correlation. We therefore model the uncertain correlation with an info-gap model and embed the hypothesis in an info-gap robustness function.

We formulate a binary hypothesis test with independent measurements in section 2. This hypothesis test can be used to decide on the level of confidence of a specific value of the probability of absence, if the measurements are known to be statistically independent. In section 3 we formulate a non-probabilistic info-gap model for uncertainty in the correlations between the measurements (Ben-Haim, 2006). The hypothesis test is then extended in section 4 to deal with the uncertain correlations between the measurements. This is achieved by introducing the robustness function for an hypothesized value of the probability of absence. The basic properties of the robustness function are discussed and illustrated by an example in section 5. The paper concludes in section 6.

## 2 Binary Hypothesis Test with Independent Measurements

Consider a sequence of statistically independent measurements where each measurement is either 'null' or 'positive'. For instance a null result is that an agent (e.g. disease) is not found, while a positive result is that the agent is found. Let $p$ denote the probability of a null result. (This can be further resolved into the probability of absence and the probability of a null given absence. We will not make this distinction since we are not modeling the measurement process itself.) We now formulate a standard binary hypothesis test (DeGroot, 1986).

Let $p_{\text{nom}}$ be an hypothesized value of $p$. We worry that the true probability of a null result is lower than $p_{\text{nom}}$ and consequently the probability of a positive (pernicious) result is greater than $1 - p_{\text{nom}}$. We have observed $n$ null results in $n$ trials, and we wish to test between the following two hypotheses:

$$H_0 : \qquad p = p_{\text{nom}} \tag{1}$$

$$H_1 : \qquad p < p_{\text{nom}} \tag{2}$$

That is, having observed $n$ nulls in $n$ trials, at what confidence can we reject the hypothesis that $p = p_{\text{nom}}$ in favor of the hypothesis that $p$ is smaller, considering only aleatoric uncertainty? Epistemic uncertainty about the degree of correlation will be introduced in section 3.

The level of significance of this test is the probability, conditioned on $H_0$, of a result which impugns $H_0$ more than the current observation ($n$ nulls in $n$ trials). $H_0$ would be weakened, with respect to $H_1$, if fewer than $n$ nulls were observed. That is, the level of significance is:

$$\alpha = \text{Prob}(m < n | H_0) \tag{3}$$

where $m$ is the number of nulls and $n$ is the number of measurements. Given the assumption that the measurements are statistically independent, the distribution of nulls is binomial and the level of significance becomes:

$$\alpha = 1 - p_{\text{nom}}^n \tag{4}$$

Inverting this we find that the smallest value of $p_{\rm nom}$ for which $H_0$ is rejected at level of significance $\alpha$ is:

$$p_{\rm nom} = (1-\alpha)^{1/n} \tag{5}$$

We will refer to this value of $p_{\rm nom}$ as the "nominal implied probability" of a null at level of significance $\alpha$, given $n$ null observations out of $n$ trials. Speaking even less precisely (but more suggestively) we will sometimes say that, given $n$ nulls out of $n$ trials, the probability of a null is no smaller than $p_{\rm nom}$ at level of significance $\alpha$.

## 3 Uncertain Correlation

Let $0_n$ denote the observation of $n$ nulls out of $n$ trials. We will henceforth consider $n > 1$ (though this assumption is not needed at all steps of the following argument). If the measurements are all statistically independent then the probability of $0_n$, conditioned on $H_0$, is $p_{\rm nom}^n$. However, if some or all of the measurements are statistically dependent, implying the existence of correlation among some of the measurements, then the probability of $0_n$ can be either greater or less than $p_{\rm nom}^n$.

If all the measurements are completely correlated, then the probability of $0_n$ equals the probability of any single measurement being null which, under $H_0$, is $p_{\rm nom}$.

At the other extreme, if two or more of the measurements are strictly anti-correlated then the probability of all measurements being null is precisely zero.

Most generally, we see that the probability of $0_n$, conditioned on $H_0$ and for $n > 1$, can take any value between 0 and $p_{\rm nom}$, depending on the degree of correlation among the measurements:

$$0 \le {\rm Prob}(0_n|H_0) \le p_{\rm nom} \tag{6}$$

The extreme values, 0 and $p_{\rm nom}$, occur for extreme correlations or anti-correlations. We do not have reason either to believe or to reject these extreme cases or any of the uncountable infinity of intermediate cases. Consequently, we formulate an info-gap model for uncertainty in the probability of $0_n$, conditioned on $H_0$, which can deviate from the nominal—uncorrelated—situation by some unknown fraction $h$:

$$\mathcal{U}(h) = \left\{ {\rm Prob}(0_n|H_0) = p_{\rm nom}^n + u : \ -p_{\rm nom}^n h \le u \le (p_{\rm nom} - p_{\rm nom}^n) h \right\}, \quad 0 \le h \le 1 \tag{7}$$

The family of sets $\mathcal{U}(h)$ becomes more inclusive as $h$ increases from 0 to 1. When $h = 0$ then $\mathcal{U}(h)$ contains only the nominal estimate, $p_{\rm nom}^n$. When $h = 1$ then $\mathcal{U}(h)$ equals the entire interval in eq.(6). The sets $\mathcal{U}(h)$ thus obey the axioms of contraction and nesting which characterize all info-gap models and which endow $h$ with its meaning as an "horizon of uncertainty".

There are many ways in which one could represent info-gaps in the correlations among observations. The choice of an info-gap model depends on the available information and on judgment by the analyst. The info-gap model of eq.(7) derives naturally from the probability model inherent in this problem.

## 4 Robustness to Uncertainty

We are unable to use the hypothesis test of eqs.(1) and (2) to estimate the probability of a null because the conditional probability, ${\rm Prob}(0_n|H_0)$, is uncertain as represented by the info-gap model of eq.(7). We can however evaluate the robustness to uncertainty of any estimate, $p_{\rm e}$, of the probability of a null. That is, we ask, What is the greatest horizon of uncertainty in the correlation up to which $p_{\rm e}$ errs no more than an acceptable amount? We now formulate this precisely.

**Formulation.** The conditional probability of $0_n$ is ${\rm Prob}(0_n|H_0)$ which is uncertain. Nonetheless, arguing as in eq.(4), the level of significance is:

$$\alpha = 1 - {\rm Prob}(0_n|H_0) \tag{8}$$

At any horizon of uncertainty $h$ this can be written:

$$\alpha \;=\; 1-(p_{\rm nom}^n+u) \tag{9}$$

where $u$ is constrained to the interval specified in $\mathcal{U}(h)$ in eq.(7). Inverting this as in eq.(5), the uncertain implied probability is:

$$p_{\rm nom}(u)=(1-\alpha-u)^{1/n} \tag{10}$$

We know neither the value of $u$ nor the value of horizon of uncertainty $h$. However, there are some values of $u$ and $h$ at which $p_{\rm nom}(u)$ in eq.(10) is the correct statistical estimate of a null result at level of significance $\alpha$.

Let $p_{\rm e}$ be any estimate of the probability of a null result. $p_{\rm e}$ may be the nominal implied probability of a null, eq.(5), or any other value. Our performance requirement is that $p_{\rm e}$ differ from the correct statistical estimate, $p_{\rm nom}(u)$ in eq.(10), by no more than $\delta$ (which is a parameter we choose to reflect the required accuracy):

$$|p_{\rm nom}(u)-p_{\rm e}|\le\delta \tag{11}$$

The robustness (to uncertainty in the degree of correlation) of estimate $p_{\rm e}$ is the greatest horizon of uncertainty $h$ at which eq.(11) is obeyed:

$$\widehat{h}(p_{\rm e},\delta)=\max\left\{h:\ \left(\max_{u\in\mathcal{U}(h)}|p_{\rm nom}(u)-p_{\rm e}|\right)\le\delta\right\} \tag{12}$$

There is a slight abuse of notation in the expression $u\in\mathcal{U}(h)$. The elements of the info-gap model are actually values of $\text{Prob}(0_n|H_0)$. This abuse should cause no confusion and is notationally simpler.

**Derivation.** Let $\mu(h)$ denote the inner maximum in eq.(12). This is a monotonically increasing function of $h$ because the sets of the info-gap model are nested according to $h$. The robustness is the greatest value of $h$ at which $\mu(h)=\delta$. Consequently, a plot of $\mu(h)$ vs. $h$ is the same as a plot of $\delta$ vs. $\widehat{h}(p_{\rm e},\delta)$. That is, $\mu(h)$ is the inverse of the robustness, $\widehat{h}(p_{\rm e},\delta)$, viewed as a function of $\delta$ at fixed $p_{\rm e}$. We will now develop an expression for $\mu(h)$.

Referring to eq.(10), we note that $|p_{\rm nom}(u)-p_{\rm e}|$ takes its maximum absolute value when $u$ takes one of its extreme values at horizon of uncertainty $h$. Referring to the info-gap model of eq.(7), we find the following expression for $\mu(h)$ as the greater between two alternatives:

$$\mu(h)=\max\left\{\left|[1-\alpha-(p_{\rm nom}-p_{\rm nom}^n)h]^{1/n}-p_{\rm e}\right|,\ \left|[1-\alpha+p_{\rm nom}^n h]^{1/n}-p_{\rm e}\right|\right\} \tag{13}$$

This is a computationally convenient expression for the inverse of the robustness function. The important aspect of this relation is that the maximum between the two terms in eq.(13) can change as $h$ increases. This entails the possibility of discontinuity in the robustness, which will lead to the intersection between robustness curves for some alternative choices of $p_{\rm e}$ as we will see shortly.

## 5 Example

We now illustrate these ideas with several examples.

Fig. 1 shows robustness curves for the nominal implied probability of a null event, $p_{\rm nom}$ in eq.(5), for three different sample sizes $n$. Recall that $n$ is the number of observed null results. Three ideas should be noted: trade-off, zeroing and the cost of robustness.

**Trade-off.** The positive slope expresses the trade-off between robustness, $\widehat{h}(p_{\rm nom},\delta)$, and error in the estimated probability of a null, $\delta$. Large robustness against uncertainty in the degree of correlation is obtained only by accepting the possibility of large error in estimating the probability of null events. Conversely, small error entails low robustness.

**Zeroing.** If there is no correlation between the measurements then the nominal implied probability, eq.(5), is the correct statistical estimate of the probability of a null event at level of significance $\alpha$. That is, in the absence of correlation, the error, $\delta$, is zero. However, arbitrarily small correlation can

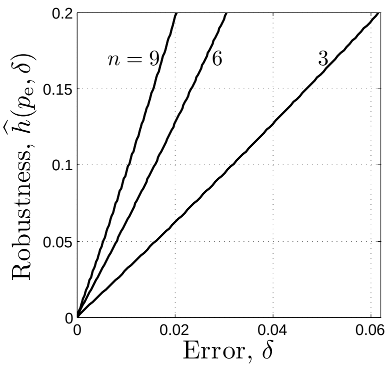


Figure 1: Robustness vs. error in probability of nulls, for 3 sample sizes. $\alpha = 0.05$, $p_e = p_{nom}$.

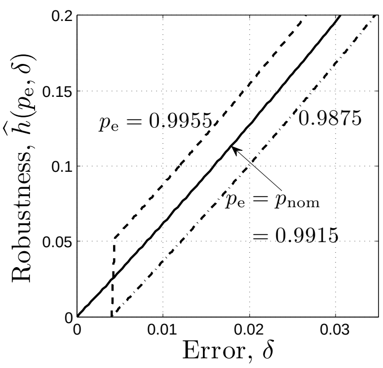


Figure 2: Robustness vs. error in probability of nulls, for various estimated probabilities. $n = 6$, $\alpha = 0.05$.

cause this estimate to err. Consequently, the robustness to error of the nominal implied probability is zero. This is manifested by the robustness curves reaching the horizontal axis (zero robustness) at zero error of the nominal implied probability.

**Cost of robustness.** The slope of the robustness curve increases with the sample size. A large slope means that a large increase of robustness can be obtained in exchange for a small increase in the error of the estimated probability. A small slope implies the opposite: the cost of robustness—in units of error—is large. We note that the cost of robustness decreases as the sample size increases, but that the utility of the marginal measurement also decreases with increasing sample size.

Fig. 2 show robustness curves for three different choices of $p_e$. The solid curve uses the nominal implied probability, $p_{nom}$, and is reproduced from fig. 1 for the case $n = 6$. The dashed curve is for a value of $p_e$ which exceeds $p_{nom}$ and the dot-dashed curve is for $p_e < p_{nom}$.

The first thing to note in fig. 2 is that the solid and dot-dash curves have smooth slope while the dashed curve has a kink. This occurs because the first two cases are determined entirely by the right-hand case in eq.(13), while the dashed curve switches from the first to the second case at the kink. To understand this in more detail, recall that $p_{nom} = (1 - \alpha)^{1/n}$. Examination of eq.(13) shows that the second term is always the maximum when $p_e \le p_{nom}$, hence the solid and dot-dash curves are smooth. However, when $p_e > p_{nom}$, then the first term in eq.(13) is greater than the second for small $h$, and the second term is greater for large $h$. The kink occurs at the transition. This kink has a very important implication, as we now see.

**Preference reversal.** The nominal error of $p_e$ is positive when $p_e$ differs from $p_{nom}$, so the dashed and dot-dashed curves intersect the horizontal axis at $\delta > 0$ (this is the zeroing property discussed earlier). These two $p_e$ values are chosen to have the same nominal error. However, when $p_e > p_{nom}$ (dashed curve) the cost of robustness is very small (very large slope) near the horizontal axis. This causes the dashed curve to rise very rapidly and to cross the solid curve. Thus, when $p_e > p_{nom}$, the robustness of $p_e$ exceeds the robustness of $p_{nom}$ over part of the range of $\delta$. Specifically, $p_{nom}$ is more robust than the larger $p_e$ for $\delta$ from zero to about 0.0044, though the robustness of $p_{nom}$ is relatively low in this range. On the other hand, for $\delta > 0.0044$ we see that the robust preference is for the larger value of $p_e$. The kink in the robustness curve entails the possibility for a reversal of preference between $p_{nom}$ and values of $p_e$ which are greater than $p_{nom}$.

Let us examine the solid and dashed curves in fig. 2 more closely at estimation error $\delta = 0.0044$, which is where the kink occurs in the dashed curve. The robustnesses of the nominal estimate, $p_{nom}$, and the dashed alternative estimate, $p_e$, are 0.025 and 0.050 respectively. Referring to the info-gap model in eq.(7) we see that $p_e$ can tolerate twice as large a range of uncertainty in the probability resulting from unknown correlations.

We can also construct robustness-based confidence intervals for $p_{\rm nom}$ and $p_{\rm e}$ at $\delta = 0.0044$:

$$p_{\rm nom}: \quad 0.9915 \pm 0.0044 = [0.9871,\ 0.9959], \quad \widehat{h}(p_{\rm nom}, \delta) = 0.025 \tag{14}$$

$$p_{\rm e}: \quad 0.9955 \pm 0.0044 = [0.9911,\ 0.9999], \quad \widehat{h}(p_{\rm e}, \delta) = 0.050 \tag{15}$$

These are intervals of values for the probability of a null event, $p$. Specifically, these intervals contain the $\alpha$-confidence intervals for all horizons of uncertainty up to the robustness. The intervals for $p_{\rm nom}$ and $p_{\rm e}$ overlap. However, the interval based on $p_{\rm e}$ contains substantially larger values, and the robustness to uncertain correlations is twice as large for $p_{\rm e}$ as for $p_{\rm nom}$. We will return to robustness-based confidence intervals after extending the discussion of curve-crossing.

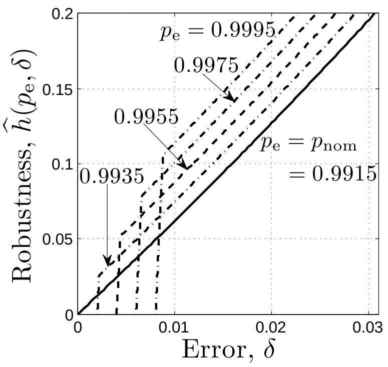


Figure 3: Robustness vs. error in probability of nulls, for various estimated probabilities. $n = 6$, $\alpha = 0.05$.

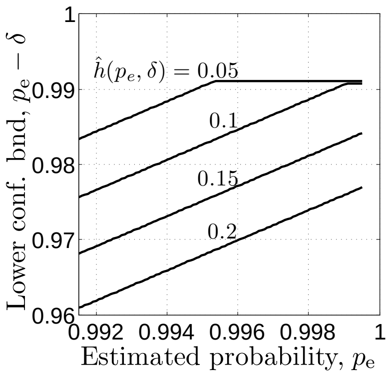


Figure 4: Lower confidence bound vs. estimated error, at various fixed robustnesses. $n = 6$, $\alpha = 0.05$.

Fig. 3 illustrates the curve-crossing phenomenon for 4 values of $p_{\rm e}$ which exceed $p_{\rm nom}$. The solid and dashed curves are the same as in fig. 2. The nominal value of implied probability is $p_{\rm nom} = 0.9915$. The figure shows $p_{\rm e}$ values of 0.9935. 0.9955, 0.9975 and 0.9995. We see that the robustness gain with $p_{\rm e}$ increases (making $p_{\rm e}$ more attractive) as $p_{\rm e}$ rises further above $p_{\rm nom}$, but the curve crossing occurs at a greater value of $\delta$ (making $p_{\rm e}$ less attractive). The analyst must make a value judgment in choosing $p_{\rm e}$, balancing between robustness, $\widehat{h}(p_{\rm e}, \delta)$, (which is large for larger $p_{\rm e}$) and estimation error $\delta$ (which is small for smaller $p_{\rm e}$).

Fig. 4 shows an alternative representation of our results, providing further support for the judgment to be made. The horizontal axis is the estimated probability of a null, $p_{\rm e}$. Any value of $p_{\rm e}$ and of robustness, $\widehat{h}(p_{\rm e}, \delta)$, induce a robustness-based confidence interval, $p_{\rm e} \pm \mu(\widehat{h})$ (truncated at 0 or 1 as needed), as explained in eqs.(14) and (15). The vertical axis in fig. 4 is the lower bound of this confidence interval. (The upper bound of the confidence interval is not shown since it is equal or very close to unity.) The slope of the curve will be either 2 or 0, depending on which term dominates in the expression for $\mu(h)$ in eq.(13).

The positive slopes in fig. 4 express the tightening of the robustness-based confidence interval as the estimated value, $p_{\rm e}$, increases. We also see that the curves shift upward as the robustness decreases, indicating the tightening of the confidence interval as the robustness decreases. This results from the trade-off between robustness, $\widehat{h}(p_{\rm e}, \delta)$, and estimation error $\delta$.

Figs. 3 and 4 together assist the analyst in choosing and evaluating an estimate of the null probability. Fig. 3 emphasizes the trade-off between estimation error and robustness to unknown correlation. The curve crossing in this figure highlights the robustness-advantage of $p_{\rm e}$ values larger than the nominal estimate, which is obtained at the expense of enlarged estimation error. Fig. 4 emphasizes this trade-off differently, as a tightening of the confidence interval as the robustness decreases.

## 6 Conclusion

Null events—not detecting a pernicious agent—are the basis for declaring the agent is absent. Repeated nulls strengthen confidence in the declaration. However, correlations between observations are difficult to assess in many situations and introduce uncertainty in interpreting repeated nulls.

We quantify uncertain correlations using an info-gap model, which is an unbounded family of nested sets of possible probabilities. An info-gap model is non-probabilistic and entails no assumption about a worst case. We then evaluate the robustness, to uncertain correlations, of estimates of the probability of a null event. This is then the basis for evaluating a robustness-based confidence interval for the probability of a null.

$p_{\text{nom}}$ denotes the nominal estimate of the probability of a null (based on assuming no correlations) at level of significance $\alpha$, eq.(5). We compare the robustness of $p_{\text{nom}}$ with other estimates of the probability of a null, $p_{\text{e}}$.

The purpose of the measurements is to determine the confidence with which absence of the pernicious agent can be declared. The analyst asks, What is the greatest value of the probability of a null which can confidently be adopted, based on $n$ nulls in $n$ observations.

$p_{\text{nom}}$ is the preferred estimate if the measurements are uncorrelated. However, if there is uncertainty in the correlations between the measurements, then the choice of an estimate is more complicated. We show the following results.

- $p_{\text{nom}}$ has greater robustness to uncertain correlations than smaller estimates of the probability of a null, $p_{\text{e}} < p_{\text{nom}}$.
- $p_{\text{nom}}$ has greater robustness to uncertain correlations than larger estimates of the probability of a null, $p_{\text{e}} > p_{\text{nom}}$, for estimation error less than an identified threshold.
- $p_{\text{nom}}$ has lower robustness to uncertain correlations than larger estimates of the probability of a null, $p_{\text{e}} > p_{\text{nom}}$, for estimation error greater than an identified threshold.

The first item indicates that considerations of robustness would never lead one to prefer a lower-than-nominal estimate over the nominal estimate of probability of absence of the pernicious agent. Thus considerations of robustness would never lead one to choose a greater-than-nominal estimate of the probability of presence of the pernicious agent.

The second and third items indicate that the preference between higher-than-nominal and nominal estimates of the probability of a null may change, depending on the estimation error which is acceptable, and the corresponding robustness to uncertain correlation.

It may at first seem counter-intuitive that the introduction of uncertainty could lead one to adopt a greater-than-nominal probability of a null, making the picture seem rosier than in the absence of uncertainty. The intuition to grasp, however, is that an estimate has two attributes: nominal error and cost of robustness. The nominal estimate, $p_{\text{nom}}$, has zero nominal error (its robustness curve intersects the axis at $\delta = 0$). Any other estimate, $p_{\text{e}}$, has greater nominal error (so its robustness curve intersects the axis at $\delta > 0$). However, the costs of robustness of $p_{\text{nom}}$ and $p_{\text{e}}$ may also differ. In particular, when $p_{\text{e}} > p_{\text{nom}}$, the initial cost of robustness is very low for $p_{\text{e}}$, causing a robustness advantage for $p_{\text{e}}$ over $p_{\text{nom}}$ at positive error ($p_{\text{e}}$'s steep robustness curve rises above $p_{\text{nom}}$'s curve). Since some error must be tolerated in order to have any robustness, this can result in preference for greater-than-nominal probability of a null. The two attributes—nominal error (intercept) and cost of robustness (slope)—combine to produce this result. The introduction of uncertain correlation does not make the picture rosier. Now the analyst must deal with both the statistical error (with confidence quantified by the level of significance $\alpha$) and the additional info-gap error (with confidence quantified by the robustness $\widehat{h}(p_{\text{e}}, \delta)$). Managing these two foci of uncertainty can lead to altered preference among potential estimates of the probability of a null.

**Acknowledgement**

The author is indebted to Mark A. Burgman, Ayala Cohen, David Fox, Malka Gorfine, Mick McCarthy, L. Joe Moffitt, Andrew Robinson, John K. Stranlund, Miriam Zacksenhouse and an anonymous reviewer for valuable comments. Funding for this research was provided by the U.S. Department of Agriculture under USDA/ERS/PREISM Cooperative Agreement No.58-7000-8-0095.

## 7 References

1. Ben-Haim, Yakov, 2006, *Info-gap Decision Theory: Decisions Under Severe Uncertainty,* 2nd edition, Academic Press, London.

2. Coolen, F.P.A., 2006, On probabilistic safety assessment in the case of zero failures. *Journal of Risk and Reliability,* vol. 220, pp.105–114.

3. DeGroot, Morris H., 1986, *Probability and Statistics,* 2nd ed., Addison-Wesley, Reading, MA.

4. Franklin, Donald C., 1999. Evidence of disarray amongst granivorous bird assemblages in the savannas of northern Australia, a region of sparse human settlement, *Biological Conservation,* vol.90, pp.53–68.

5. Knight, Frank H., 1921, *Risk, Uncertainty and Profit,* Hart, Schaffner and Marx. Re-issued by Harper Torchbooks, New York, 1965.

6. Lueck, Dean and Jeffrey A. Michael, 2003, Preemptive Habitat Destruction Under The Endangered Species Act, *Journal of Law and Economics,* vol.XLVI, pp.27–60.

7. Lyons, Melinda, Sally Adams, Maria Woloshynowych and Charles Vincent, 2004, Human reliability analysis in healthcare: A review of techniques, *International Journal of Risk & Safety in Medicine,* vol.16, no.4, pp.223–237.

8. Moffitt, L.Joe, John K.Stranlund, and Barry C.Field, 2005, Inspections to Avert Terrorism: Robustness Under Severe Uncertainty, *Journal of Homeland Security and Emergency Management,* vol.2, no. 3. http://www.bepress.com/jhsem/vol2/iss3/3.

9. Moffitt, L.Joe, John K.Stranlund and Craig D.Osteen, 2008, Robust detection protocols for uncertain introductions of invasive species, *Journal of Environmental Management,* vol.89, pp.293-299.

10. Regan, Tracey J., Michael A.McCarthy, Peter W.J.Baxter, F.Dane Panetta and Hugh P.Possingham, 2006, Optimal eradication: when to stop looking for an invasive plant, *Ecology Letters,* vol.9: pp.759-766.

11. Rout, T.M., C.J.Thompson, and M.A.McCarthy, 2009, Robust decisions for declaring eradication of invasive species, *Journal of Applied Ecology,* vol.46, pp.782-786.

12. Schweppe, F.C., 1973, *Uncertain Dynamic Systems,* Prentice-Hall, Englewood Cliffs.

13. Vincent, Charles, 2006, *Patient Safety,* Elsevier, Amsterdam.

14. Wald, Abraham, 1945, Statistical decision functions which minimize the maximum risk, *Annals of Mathematics,* vol. 46, no. 2, pp.265–280.