Bryce Terwilliger
Advisor: Ian Abramson
June 2, 2011

Contents

1 Abstract

In this paper we consider an a randomized test for a shift in a non-parametric setting developed by Bell and Docsum and by an alternative means find the asymptotic relative efficiency is 1.

2 Introduction

Given two independent samples $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ from populations with continuous cumulative distribution functions $F_0(x)$ and $F_\delta(x) = F_0(x-\delta)$. We will consider testing the null hypothesis:

$$H_0 : \delta = 0$$

against the one-sided alternative hypothesis

$$H_1 : \delta > 0.$$

This is a classic set up for the t-test or z-test in the case where

$$F_0(x) = \Phi(x)$$

where Φ is the cumulative density of a normal distribution. If it is suspected that F_0 is not normal then it is reasonable to consider a non-parametric test such as the Mann-Whitney U test. The asymptotic cost of using the Mann-Whitney test over the classic t-test when the F_0 really is normal is given by asymptotic relative efficiency $\frac{3}{\pi}$.

2.1 Asymptotic Relative Efficiency

Asymptotic relative efficiency is a means of determining the power of one test against another with large sample sizes. In this case, we will consider a sequence of pairs of populations that tend to the null at a $\sqrt{n}$ rate. On this sequence of alternatives we compare the sample size of each test required to attain fixed α and β levels such that the power is between 0 and 1. If the limit of the ratio of the sample sizes exists then that ratio is the ARE. More technically, given a sequence of estimators, δ_n of $g(\theta)$ satisfying

$$\sqrt{n}[\delta_n - g(\theta)] \to N(0, \tau^2)$$

and a sequence of estimators $\delta'_{n'}$, where $\delta'_{n'}$ is based on $n' = n'(n)$ observations, also satisfies $\sqrt{n}[\delta'_{n'} - g(\theta)] \to N(0, \tau^2)$, then the asymptotic relative efficiency of $\{\delta_n\}$ with respect to $\{\delta'_{n'}\}$ is

$$\lim_{n\to\infty} \frac{n'(n)}{n},$$

provided the limit exists and is independent of the subsequences n'.

To recover this loss of efficiency we will consider a randomized test developed by Bell and Doksum [1]. In this test, an observation of rank i in the pooled original data will be replaced by an observation of rank i in an independent normal sample. The difference of the means of the new samples is the statistic we will consider.

Under the null $F_0(x) = F_\delta(x)$, the probability that rank(x_i), in the pooled sample, is less than rank(y_j) is .5 because F_0 is continuous for all i and j. So, the sample that replaces the $x's$ is iid standard normal, as is the case for the $y's$. Thus, the z-test is a justified test for determining a difference in the means of the replacement samples. This is a direct computation done in [1].

Lemma 1. *Let F be a continuous cpf and let H be any cpf. If $W_1, W_2, \ldots, W_N$, and $Z_1, Z_2, \ldots, Z_N$ are independent random samples with cpf's F and H, respectively, if $R(W_i)$ denotes the rank of W_i among $W_1, W_2, \ldots, W_N$, and if $Z(i)$ is the i^{th} order statistic of $Z_1, Z_2, \ldots, Z_N$; then $Z(R(W_1)), Z(R(W_2)), \ldots, Z(R(W_N))$ have the same joint distribution as the random sample of $Z_1, Z_2, \ldots, Z_N$.*

Proof. Let A_N be a Borel set in N dimensional Euclidean space.

$P\{Z(R(W_1)), \ldots, Z(R(W_N))] \in A_N\}$
$= \sum P\{[Z(r_1), \ldots, Z(r_N)] \in A_N | R(W_1) = r_1, \ldots, R(W_N) = r_N\} P\{R(W_1) = r_1, \ldots, R(W_N) = r_N\}$
where the sum is over all the possible permutations of $\{r_1, \ldots, r_N\}$ of the ranks$\{1, \ldots, N\}$
$= \dfrac{\sum P\{[Z(r_1), \ldots, Z(r_N)] \in A_N\}}{N!}$
$= \sum P\{[Z(r_1), \ldots, Z(r_N)] \in A_N | R(Z_1) = r_1, \ldots, R(Z_N) = r_N\} P\{R(Z_1) = r_1, \ldots, R(Z_N) = r_N\}$
$= P\{[Z_1, \ldots, Z_n]\} \in A_N\}$

Since $P\{Z(R(W_1)), \ldots, Z(R(W_N))] \in A_N\} = P\{[Z_1, \ldots, Z_n\} \in A_N\}$ for each Borel set, $Z(R(W_1)), Z(R(W_2)), \ldots, Z(R(W_N))$ have the same join distribution as the random sample of $Z_1, \ldots, Z_N$. □

Under the alternative distribution, the thinned sample does not have a normal distribution. In fact for a fixed sample size the pseudo $X's$ may not be independent from the pseudo $Y's$ or indeed from each other although they are exchangeable because the pseudo $X's$ and $Y's$ have been reordered. This is a problem because we would like to use the central limit theorem to compute the ARE of the difference.

2.2 A poissonization approach

Poissonization is a device we used to create the random samples because the number of one size interval will be independent of the one in a disjoint interval. Fix n, (it will tend to ∞ later).

Let $\mathbb{X}$ be a non-homogeneous poisson point process on $\mathbb{R}$, following intensity function

$$\lambda_{n,0}(x) = n\phi(x)$$

This will scatter a poisson number:

$$M = M_n \sim \text{pois}(n) \text{ of points } x_1, ..., x_M \text{ on } \mathbb{R}$$

Fixing M, (i.e. conditioning on it), the x's were an iid sample of size M from $\mathcal{N}(0, 1)$. (Needs a theorem to justify)

Independently let $\mathbb{Y}$ be a non-homogeneous poisson point process following intensity function

$$\lambda_{n,\delta}(y) = n\phi(y - \delta)$$

(δ puts us on the alternative; later we'll take a sequence of δ's that tend to 0 at some rate.) This will also scatter a poisson number

$$N = N(n) \text{ of points } y_1, ..., y_N$$

(As if conditioning on N, the Y's were an iid sample of size N from $\mathcal{N}(\delta, 1)$. (Needs the same theorem to justify)

Now superposing the two processes we have a poisson point process. $\mathcal{W}$ say, with intensity: $n(\phi(w) + \phi(w - \delta))$, of which $\mathbb{X}$ and $\mathbb{Y}$ are thinned versions, hence poisson processes in their own right.

2.3 We can extend to different sample sizes

We can extend to a more general asymptotic setting, where the intensity function for $\mathbb{Y}$ is $kn\phi(y - \delta)$ for some fixed $k > 0$, allowing for the sample sizes to grow in some asymptotic proportion to each other. Extension to proportional growth rate is a useful generalization that is not hard, but it is not done here.

2.4 Conditioning on M and N

Now, conditioning on M and N (or just their sum), but otherwise independently of the mechanism that produced the $X's$ and the $Y's$. Let:

$$z_1, ..., z_{M+N} \sim \text{iid } \mathcal{N}(0, 1)$$

put the order statistics $W_{(i)}$ into order correspondence with the $Z_{(i)} : i = 1, ..., M + N$ and by referring back to the sample identity labels X or Y, that were pooled to form the $W's$, pull out an ordered subsample of M pseudo-$X's$, $x^*_{(1)}, ..., x^*_M$ where for each i

$$x^*_{(i)} \text{ is the } i^{th} \text{ largest among the ordered } Z's \text{ that "come from an } X\text{"}$$

Do the corresponding thing to get an ordered subsample of N pseudo $Y's$, $\{Y^*_{(i)}\}$.

Give the $\{X^*_{(i)}\}$ and $\{Y^*_{(i)}\}$ some uniform random shuffles. (ie equal probability for all permutations). These shuffles should be independent of each other, and everything else that has gone before, except the M and N, which are conditioned on.

$$\text{This will give unordered sets } \{X^*_i\} \text{ and } \{Y^*_i\}.$$

With M, N conditioned on these are in fact iid independent samples from two densities. What are they?

Well, let's recognize too that these ensembles are realizations of two new nonhomogeneous poisson processes. We need to identify $\lambda's$ the intensity functions; they are not necessarily even normal in shape. To get around this problem we will use a linear perturbation formula for the inverse functions.

3 A Linear Perturbation Formula for Inverse Functions

3.1 Set Up

Define the symbol $o_w(\delta)$:

$$o_w(\delta) \in \{f(\delta, w) : \mathbb{R}^2 \to \mathbb{R} | \lim_{\delta \to 0} \frac{f(\delta, w)}{\delta} = 0 \text{ for each } w\}$$

We will often not include the subscript w when it is clear from context that $o_w(\delta)$ is continuous in w. Let $h_0(w)$ and $g(w)$ have continuous derivatives on a closed interval $I \subset \mathbb{R}$.
Let $h_0(w)$ be have a strictly positive derivative; hence $h_0(w)$ is invertible.
Let $\{h_\delta(w) : \delta \in \mathbb{R}_{\geq 0}\}$ be a smoothly indexed family of functions. (i.e. partial derivatives in δ (for fixed w) are continuous for δ in some neighborhood of 0)
satisfying

$$h_\delta(w) = h_0(w) + \delta g(w) + o(\delta)$$

Proposition 2. *If $g : \mathbb{R} \to \mathbb{R}$ is invertible and $a, b \in \mathbb{R}$ with $b \neq 0$ then $h(x) = a + bg(x)$ is invertible with*

$$h^{-1}(y) = g^{-1}\left(\frac{y-a}{b}\right)$$

Proof.

$$\begin{aligned} h(h^{-1}(y)) =& h\left(g^{-1}\left(\frac{y-a}{b}\right)\right) \\ =& a + bg\left(g^{-1}\left(\frac{y-a}{b}\right)\right) \\ =& a + b\left(\frac{y-a}{b}\right) \\ =& y \end{aligned}$$

and the other direction:

$$\begin{aligned} h^{-1}(h(x)) =& h^{-1}(a + bg(x)) \\ =& g^{-1}\left(\frac{a + bg(x) - a}{b}\right) \\ =& g^{-1}(g(x)) \\ =& x \end{aligned}$$

□

Proposition 3. *If $o(x)$ is a contractive map, $o(0) = 0$ is the fixed point, and $f(x) = x + o(x)$ is invertible then*

$$f^{-1}(y) = y + \tilde{o}(y)$$

where $\tilde{o}(y) = -o(y - o(y - o(y - \cdots)))$. In particular, if $o(x)$ is little o in x then $\tilde{o}(y)$ is little o in y.

Proof. Fix y. $\tilde{o}(y) + y$ is the fixed point of the contractive map $x \mapsto y - o(x)$. So, the contractive mapping theorem implies $\tilde{o}(y)$ is a function of y, it does not depend on x. Then we can find the inverse of f.

$$\begin{aligned} x &= y + o(y) \\ x - o(y) &= y \\ x - o(x - o(y)) &= y \\ x - o(x - o(x - o(x - \cdots o(y)))) &= y \\ x + \tilde{o}(x) &= y \end{aligned}$$

$\tilde{o}$ a contractive map.

$$\begin{aligned} |\tilde{o}(x) - \tilde{o}(y)| =& |\tilde{o}(x) - \tilde{o}(y)| \\ =& |\tilde{o}(x) - \tilde{o}(y) - (x - y) + (x - y)| \\ =& |f^{-1}(x) - f^{-1}(y) - (x - y)| \\ =& |f^{-1}(f(\tilde{x})) - f^{-1}(f(\tilde{y})) - (f(\tilde{x}) - f(\tilde{y}))| \\ =& |\tilde{x} - \tilde{y} - (\tilde{x} + o(\tilde{x}) - \tilde{y} - o(\tilde{y}))| \\ =& | - o(\tilde{x}) + o(\tilde{y})| \end{aligned}$$

We also have $\tilde{o}(0) = 0$ since $0 = f(0) = f^{-1}(0) = 0 + \tilde{o}(0)$. □

Lemma 4. *Let functions $h_0(w)$ and $g(w)$ be C^1 functions on the closure of some bounded, open interval I and let h_0 be invertible on I. Define $h_\delta(w) = h_0(w) + \delta g(w) + o(\delta)$ where $o(\delta)$ is contractive with $o(0) = 0$ then, for small δ and $u \in h_0(I)$,*

$$h_\delta^{-1}(u) = h_0^{-1}(u) - \delta(\frac{\partial}{\partial u} h_0^{-1}(u)) g(h_0^{-1}(u)) + \tilde{o}(\delta)$$

Proof. let I^c be the bounded closure of I. Suppose with out loss of generality that h_0 is strictly increasing on I. $\frac{\partial h_0(w) + \delta g(w)}{\partial w} > 0$ for all $w \in I$ because $\frac{\partial h_0(w)}{\partial w} > \epsilon$ because h_0 is C^1 and strictly increasing on I and $g(w)$ is bounded on I^c. Then Proposition 2 gives h_δ is invertible on $h_0(I)$.

Proposition 3 states the inverse of $f(\delta) = \delta + o(\delta)$ is $f^{-1}(\delta) = \delta + \hat{o}(\delta)$. Fix u, when $g(u) \neq 0$ Proposition 2

$$h_\delta^{-1}(u) = \left(\frac{\delta - h_0(u)}{g(u)}\right) + \hat{o}\left(\frac{\delta - h_0(u)}{g(u)}\right)$$

We have $h_0^{-1}(u) = \left(\frac{0-h_0(u)}{g(u)}\right) + \hat{o}\left(\frac{0-h_0(u)}{g(u)}\right)$. Let $\tilde{o}(\delta) = \hat{o}\left(\frac{\delta-h_0(u)}{g(u)}\right) - \hat{o}\left(\frac{0-h_0(u)}{g(u)}\right)$. $\lim_{\delta\to 0}\frac{\tilde{o}(\delta)}{\delta} = 0$ because $\hat{o}$ is contractive with $\hat{o}(0) = 0$.

Finally we check

$$
\begin{aligned}
&(\frac{\partial}{\partial u} h_0^{-1}(u)) g(h_0^{-1}(u)) g(u) \\
=&\frac{g(h_0^{-1}(u))g(u)}{h_0'(h_0^{-1}(u))} \\
=&\frac{g(w)g(u)}{h_0'(w)} \\
=&\frac{g(w)g(h_0^{-1}(w))}{h_0'(w)} \\
=&1
\end{aligned}
$$

□

Theorem 5. *Let functions $h_0(w)$ and $g(w)$ be $C_0^1 : \mathbb{R} \to \mathbb{R}$ such that $h_0(u)$ is strictly increasing, $h_0^{-1}(u)$ exists and $g'(w)$ is bounded. Define $h_\delta(w) = h_0(w) + \delta g(w) + o(\delta)$. Then there exists δ_0 such that for $\delta_0 \geq \delta > 0$, h_δ^{-1} exists on a bounded interval and*

$$h_\delta^{-1}(u) = h_0^{-1}(u) - \delta(\frac{\partial}{\partial u} h_0^{-1}(u)) g(h_0^{-1}(u)) + \tilde{o}(\delta)$$

Proof. $h_0(w)$ is strictly increasing on the interval so there exists w_0 such that for all w on the interval $h_0'(w) \geq h_0'(w_0) > 0$. Let $\epsilon = h_0'(w_0)$. Since $g'(w)$ is bounded there exists δ_1 such that for all w $\delta_1 \geq \delta > 0$ $\frac{\epsilon}{2} > |\delta g'(w)|$. Since $\lim_{\delta\to 0} o(\delta) = 0$ there exists $\delta_2 > 0$ such that for $\delta_2 \geq \delta > 0$ $|o(\delta)| < \frac{\epsilon}{2}$. Then we may take $\delta_0 = \min[\delta_1, \delta_2]$ so that for $\delta_0 \geq \delta > 0$ $h_\delta(w)$ is also strictly increasing and therefore $h_\delta^{-1}(u)$ exists.

Consider the following set of graphs generated by Mathematica using the code:

```
h0[x_] := (x + .6)^6;
h01[y_] := z /. Solve[h0[z] == y, z][[6]];
h0d[x_] :=
  h0[x] + (FullSimplify[
      Normal[Series[Sin[15*z + 1], {z, 0, 11}]] /. z -> x]/50);
h01d[y_] := z /. Solve[h0d[z] == y, z][[6]];
h01d2[y_] :=
 z /. Solve[h0d[z] == y, z][[8]]; Plot[{h0[x], h01[x], h0d[x],
  h01d[x], h01d2[x], x}, {x, -.1, .35},
 PlotStyle -> {Red, Red, Blue, Blue, Blue, Dashed}, AspectRatio -> 1,
 PlotRange -> {-.1, 0.35},
 Epilog -> {Text["\!\(\*
StyleBox[SubscriptBox[
```

```
StyleBox[\"h\",\nFontSize->16], \"0\"],\nFontSize->16]\)", {.02, \
.04}], Text["\!\(\*
StyleBox[SubsuperscriptBox[
StyleBox[\"h\",\nFontSize->16], \"0\", \"1\"],\nFontSize->16]\)", \
{.04, .02}], Text["\!\(\*
StyleBox[SubsuperscriptBox[
StyleBox[\"h\",\nFontSize->16], \"\[Delta]\", \"1\"],\n\
FontSize->16]\)", {.12, .03}], Text["\!\(\*
StyleBox[SubsuperscriptBox[
StyleBox[\"h\",\nFontSize->16], \"\[Delta]\", \"1\"],\n\
FontSize->16]\)", {.036, .12}]}]
```

We will consider a small portion of this picture to illustrate the proof.

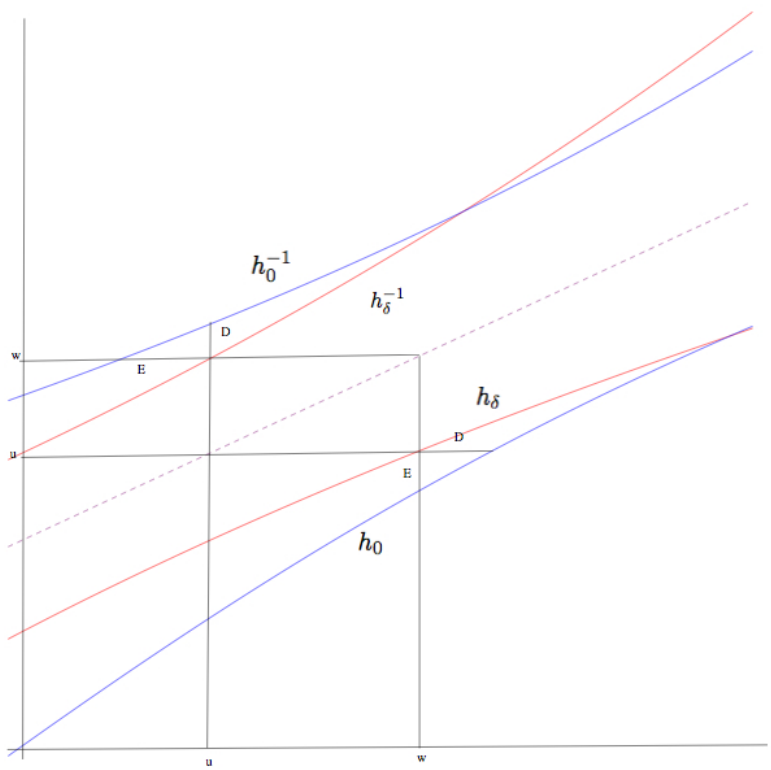

Figure 2: Diagram

Since h_0 and h_δ are C_0^1 functions we can approximate each with a linear function to arbitrary accuracy on a sufficiently small region. Fix u for all $\epsilon > 0$

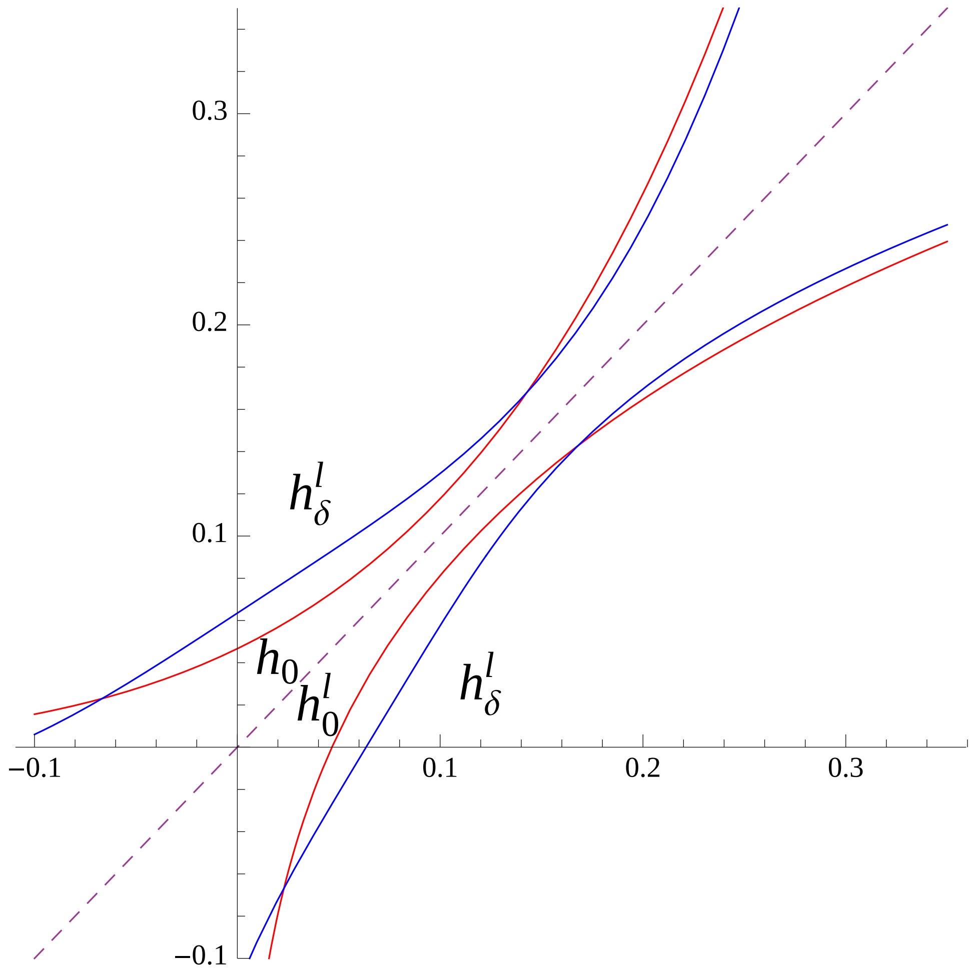

Figure 1: A large view

there exists $\delta_0 > 0$ such that $\delta_0 \geq \delta > 0$ implies

$$\max[\sup_{x,y\in[u-\delta,u+\delta]} |h_0(x) - h_0(y)|, \sup_{x,y\in[u-\delta,u+\delta]} |h_\delta(x) - h_\delta(y)|, D, E] < \epsilon$$

Without loss of generality we can assume that $h_0^{-1}(w) > h_\delta^{-1}(w)$ then

$$h_\delta^{-1}(w) = h_0^{-1}(w) - D$$

and $\frac{E}{D} \approx h_0'(w)$ implies $D \approx \frac{E}{h_0'(w)}$. Since $o(\delta) \to 0$ as $\delta \to 0$ we have

$$h_\delta(w) - h_0(w) = E \approx \delta g(w) \text{ for small } \delta$$

Then $D \approx \frac{\delta g(w)}{h'(w)}$ but we can express w in terms of u: $w = h_0^{-1}(u)$ then after a substitution

$$h_\delta^{-1}(u) = h_0^{-1}(u) - \delta(\frac{\partial}{\partial u} h_0^{-1}(u))g(h_0^{-1}(u)) + \tilde{o}(\delta)$$

where $\tilde{o}(\delta) \to 0$ as $\delta \to 0$. □

Then show that

1. $E_\delta X_1^* = -\frac{\delta}{2}$
2. $E_\delta Y_1^* = \frac{\delta}{2}$

Therefor separation is still δ and the asymptotic power will be the same for $\delta \to 0$ inversely to $\frac{1}{\sqrt{n}}$.

and Its inverse.pdf

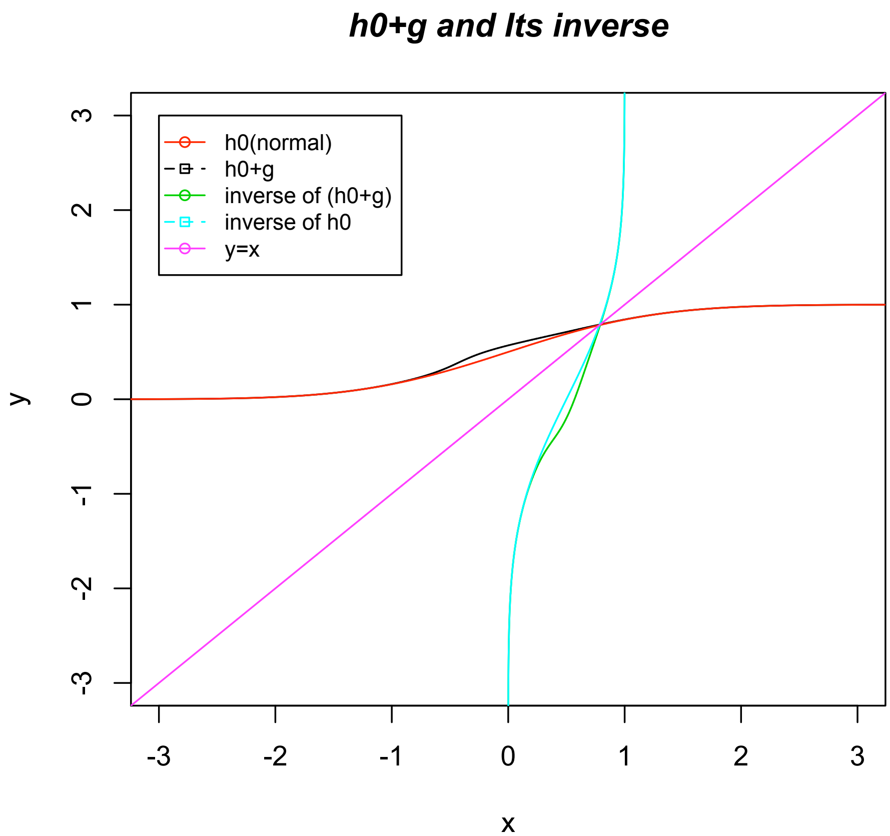

Figure 3: when $\delta = 1$ h_δ looks like

4 Computing the Mean and Variance

We need the mean and variance of x_i^*(and y_i^*):

$$\text{Mean: } E_\delta(x_1^*) = 2\int_{-\infty}^{\infty} x \frac{\phi(H_\delta^{-1}(\Phi(x)))\phi(x)}{\phi(H_\delta^{-1}(\Phi(x))) + \phi(H_\delta^{-1}(\Phi(x)) - \delta)}\, dx$$

Let:

$$H_\delta(w) = \frac{1}{2}(\Phi(w) + \Phi(w-\delta)).$$

$$\frac{\partial}{\partial\delta} H_\delta(w)\bigg|_{\delta=0} = -\frac{1}{2}\phi(w-\delta)\bigg|_{\delta=0} = -\frac{1}{2}\phi(w)$$

$$\text{then } H_\delta(w) = \Phi(w) - \frac{1}{2}\phi(w)\delta \text{ for } \delta \text{ sufficiently small}$$

We use Theorem 4 to compute the inverse. We need h_0 and g.

$$h_0(w) = \Phi(w)$$
$$g(w) = \frac{1}{2}\phi(w)$$

Then these are substituted into the formula given in Theorem 4 for the inverse.

$$H_\delta^{-1}(u) = h_0^{-1}(u) + \delta(\frac{\partial}{\partial u}h_0^{-1}(u))g(h_0^{-1}(u)) + \tilde{o}(\delta)$$
$$H_\delta^{-1}(u) = \Phi^{-1}(u) + \delta(\frac{\partial}{\partial u}\Phi^{-1}(u))\frac{1}{2}\phi(\Phi^{-1}(u)) + \tilde{o}(\delta)$$
$$H_\delta^{-1}(u) = \Phi^{-1}(u) + \delta(\frac{1}{\phi(\Phi^{-1}(u)}))\frac{1}{2}\phi(\Phi^{-1}(u)) + \tilde{o}(\delta)$$
$$H_\delta^{-1}(u) = \Phi^{-1}(u) + \frac{\delta}{2} + \tilde{o}(\delta)$$

so we have

$$H_\delta^{-1}(\Phi(x)) = x + \frac{\delta}{2} + \tilde{o}(\delta)$$

Then substituting back into the integral and omitting the error term:

$$2\int_{-\infty}^{\infty} x\frac{\phi(H_\delta^{-1}(\Phi(x)))\phi(x)}{\phi(H_\delta^{-1}(\Phi(x))) + \phi(H_\delta^{-1}(\Phi(x)) - \delta)}\,dx$$
$$= 2\int_{-\infty}^{\infty} x\frac{\phi(x+\frac{\delta}{2})\phi(x)}{\phi(x+\frac{\delta}{2}) + \phi((x+\frac{\delta}{2}) - \delta)}\,dx$$
$$= 2\int_{-\infty}^{\infty} x\frac{\phi(x+\frac{\delta}{2})\phi(x)}{\phi(x+\frac{\delta}{2}) + \phi(x-\frac{\delta}{2})}\,dx$$

Then we can also Taylor $\phi(x+\frac{\delta}{2})$ in about $\delta = 0$

$$\phi(x+\frac{\delta}{2}) = \phi(x) + \frac{\delta}{2}\phi'(x) + o(\delta)$$
$$= \phi(x)(1 - \frac{\delta x}{2}) + o(\delta)$$

similarly

$$\phi(x+\frac{\delta}{2}) = \phi(x)(1 + \frac{\delta x}{2}) + o(\delta)$$

Then substituting these two expansions back into the integral we have:

$$
\begin{aligned}
2\int_{-\infty}^{\infty} x\frac{\phi(x+\frac{\delta}{2})\phi(x)}{\phi(x+\frac{\delta}{2})+\phi(x-\frac{\delta}{2})}\,dx \\
&= 2\int_{-\infty}^{\infty} x\frac{\phi(x)(1-\frac{\delta x}{2})\phi(x)}{\phi(x)(1-\frac{\delta x}{2})+\phi(x)(1+\frac{\delta x}{2})}\,dx \\
&= 2\int_{-\infty}^{\infty} x\frac{\phi(x)(1-\frac{\delta x}{2})}{(1-\frac{\delta x}{2})+(1+\frac{\delta x}{2})}\,dx \\
&= \int_{-\infty}^{\infty} x\phi(x)(1-\frac{\delta x}{2})\,dx \qquad (\sharp) \\
&= \int_{-\infty}^{\infty} x\phi(x)\,dx - \frac{\delta}{2}\int_{-\infty}^{\infty} x^2\phi(x)\,dx \\
&= -\frac{\delta}{2}
\end{aligned}
$$

Thus we have $E_\delta(x_1^*) = -\frac{\delta}{2}$ and similarly we find $E_\delta(y_1^*) = \frac{\delta}{2}$ under the null. Starting at $\sharp$ we will compute $E_\delta({x_1^*}^2)$

$$
\begin{aligned}
E_\delta({x_1^*}^2) &= \int_{-\infty}^{\infty} x^2\phi(x)(1-\frac{\delta x}{2})\,dx \\
&= \int_{-\infty}^{\infty} x^2\phi(x)\,dx - \frac{\delta}{2}\int_{-\infty}^{\infty} x^3\phi(x)\,dx \\
&= 1
\end{aligned}
$$

Thus we have $E_\delta({x_1^*}^2) = 1 - o(\delta)$ and similarly we find $E_\delta({y_1^*}^2) = 1 - o(\delta)$ under the null.

For δ near zero x^* is very near normally distributed.

of sx.pdf

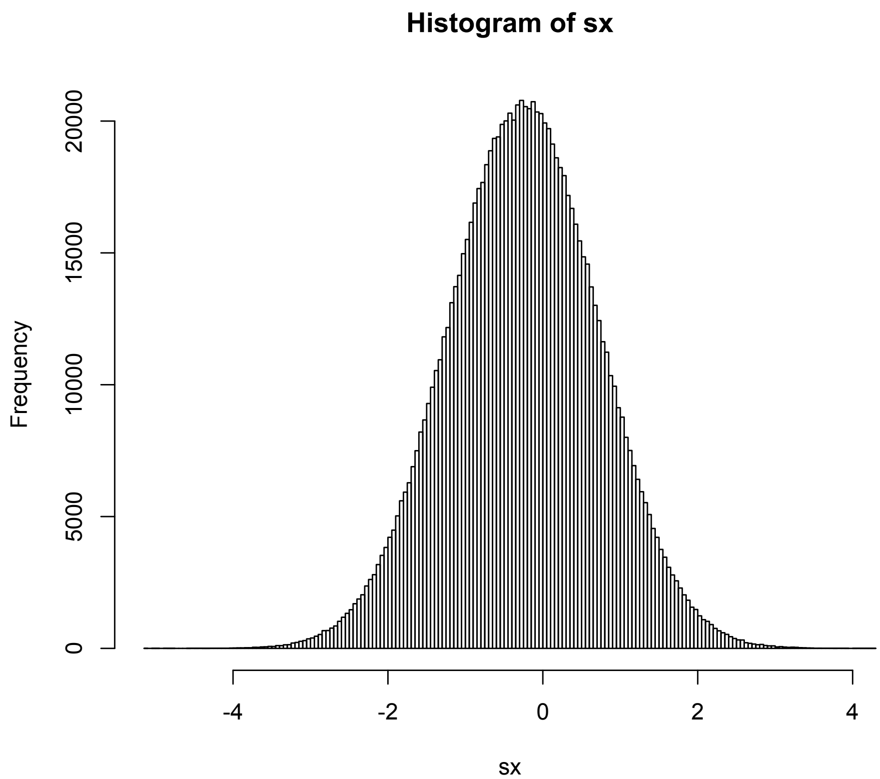

Figure 4: small δ

For large δ the pseudo x's are skewed away from zero.

of sx big delta.pdf

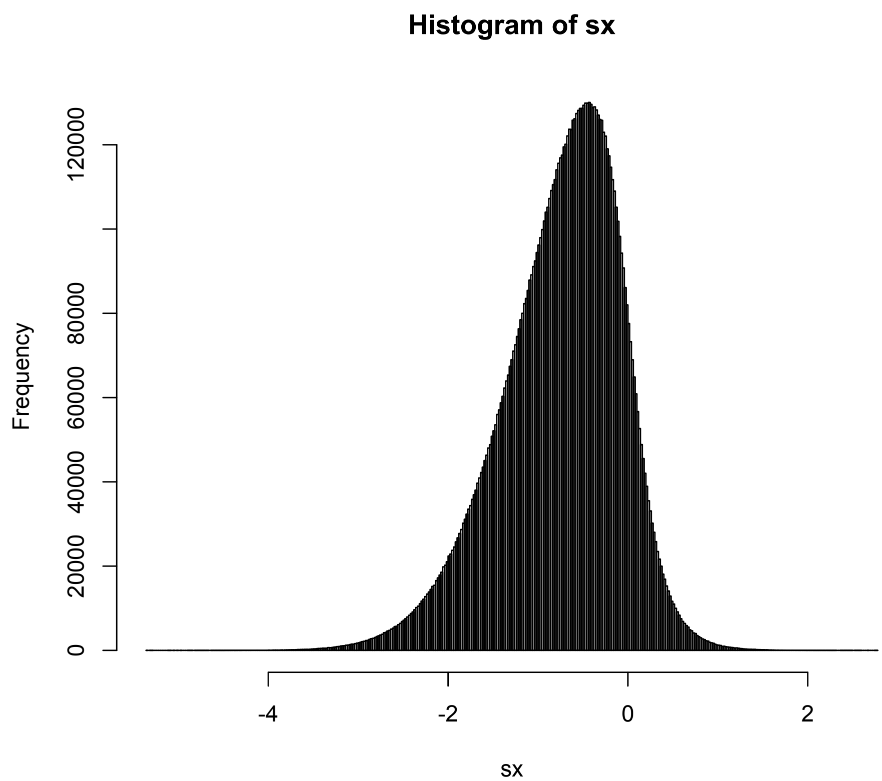

Figure 5: large δ

4.1 An application of CLT

Now we compute the distribution of the difference in the means. Note here that x_i^* and y_j^* are independent because they were generated by a thinned poisson process.

$$\lim_{n\to\infty} \sqrt{n} \left(\frac{\bar{y}^* - \bar{x}^* - \delta}{\sqrt{1 - o(\delta)}} \right) = \lim_{n\to\infty} \sqrt{n} \left(\bar{y}^* - \bar{x}^* - \delta \right) = \mathcal{N}(0,1)$$

by Slutsky's Theorem and the Central Limit Theorem. But we also have

$$\lim_{n\to\infty} \sqrt{n} \left(\bar{u} - \bar{v} - \delta \right) = \mathcal{N}(0,1)$$

where u_i are iid from $\Phi(x)$ and v_i are iid from $\Phi(x - \delta)$. Thus, the ARE of the Bell-Doksum procedure against the standard z-test, and t-test, is 1.

4.2 De-poissonization

If I had more time I would go on to undo the poisson process that generated the data to show that the data could have come from sampling a population.

5 Empirical Evidence

The following in an implementation, in R, of the test described in Bell and Doksum [1]. Various versions of this test are used in computing figures that follow.

```
RandTtest<-function(X, Y, alternative ="two.sided", paired = FALSE,
var.equal = TRUE, conf.level = 0.95){
Cx=complex(real =X, imaginary = rep(1,length(X)))
#The x's are identified with a complex value of 1 where y's have 0
XandY=sort(c(Cx,Y))
model=sort( rnorm( length( XandY)))
newX=rep(NA,length(X))
newY=rep(NA,length(Y))
k=1
j=1
for(i in 1:length(XandY)){
if(Im(XandY[i]) == 1){
newX[k] = model[i]
k=k+1}
if(Im(XandY[i])==0){
newY[j]=model[i]
j=j+1
}}
t.test(newX,newY, alternative=alternative, paired=paired,
var.equal=var.equal,conf.level=conf.level)
}
```

The following describes the ratio of sample sizes, starting at 5 and taking steps of 5 to 200, required for the Bell Doksum procedure against the one sided z-test with $\alpha = .05$ and $\beta = .2$.

[1]0.6250000	0.7142857	0.8333333	0.8333333	0.8928571	0.8823529	0.8974359	0.9302326
0.9183673	0.9615385	0.9322034	0.9230769	0.9285714	0.9589041	0.9259259	0.9302326
[17]0.9550562	0.9677419	0.9500000	0.9345794	0.9459459	0.9565217	0.9583333	0.9523810
0.9541985	0.9629630	0.9642857	0.9655172	0.9863946	0.9615385	0.9687500	0.9638554
[33]1.0000000	1.0000000	0.9668508	0.9890110	0.9840426	0.9844560	0.9653465	0.9615385

5.1 Comparison to Mann-Whitney U

It is interesting to compare the Bell Doksum procedure to the Mann-Whitney U test. I found that the U test is more powerful for sample sizes smaller than

40 and because of the ARE the Bell Doksum procedure is more powerful for large sample sizes.

```
> iteration=10000;randzSize=47;zSize=45;
>   power=0;
>  for(i in 1:iterations){
+  x=rnorm(randzSize);
+   sx=rep(NA,randzSize);
+   cx=complex(real=x,imaginary=rep(1,randzSize));
+   y=rnorm(randzSize,mean=sqrt(2)*(qnorm(.8)+qnorm(.95))/sqrt(zSize));
+   sy=rep(NA,randzSize);
+   xANDy=sort(c(cx,y));
+   z=sort(rnorm(2*randzSize));
+   a=1;b=1;
+   for(j1 in 1:(2*randzSize)){
+   if(Im(xANDy[j1])==1){sx[a]=z[j1];a=a+1}
+   else{sy[b]=z[j1];b=b+1}};
+   power=power+(qnorm(.95)<=((mean(sy)-mean(sx))*sqrt(randzSize/2)))/iterations};power
[1] 0.8017
>  iterations=10000;USize=47;zSize=45;
>  power=0;
>   for(i in 1:iterations){
+   x=rnorm(USize);
+   y=rnorm(USize,mean=sqrt(2)*(qnorm(.8)+qnorm(.95))/sqrt(zSize));
+   power=power+(.05>=(wilcox.test(x,y,alternative="less")$p.value))/iterations};power
[1] 0.8004
```

The following is a comparison of the Bell Doksum procedure against the Man Whitney U test on various sample populations.

x	y	sample size	$\begin{pmatrix}\text{power of Bell Doksum}\\ \text{power of Man Whitney U}\end{pmatrix}$
$N(0,1)$	$N(0.828825,1)$	18	$\begin{pmatrix}0.7498\\0.7582\end{pmatrix}$
$N(0,1)$	$N(0.4539661,1)$	60	$\begin{pmatrix}0.7795\\0.7787\end{pmatrix}$
$N(0,1)$	$N(0.2486475,1)$	200	$\begin{pmatrix}0.7945\\0.781\end{pmatrix}$
$exp(1)$	$exp(.5)$	18	$\begin{pmatrix}0.5289\\0.5342\end{pmatrix}$
$exp(1)$	$exp(.6)$	60	$\begin{pmatrix}0.78648\\0.77355\end{pmatrix}$
$exp(1)$	$exp(.8)$	200	$\begin{pmatrix}0.63485\\0.61019\end{pmatrix}$
$N(-.5,1)+N(.5,1)$	$N(0,1)+N(1,1)$	18	$\begin{pmatrix}0.63039\\0.64357\end{pmatrix}$
$N(-.3,1)+N(.7,1)$	$N(0,1)+N(1,1)$	60	$\begin{pmatrix}0.73273\\0.72749\end{pmatrix}$
$N(-.1.5,1)+N(.85,1)$	$N(0,1)+N(1,1)$	200	$\begin{pmatrix}0.6787\\0.66639\end{pmatrix}$
$cauchy(0,1)$	$cauchy(.2,1)$	1000	$\begin{pmatrix}0.6657\\0.793\end{pmatrix}$
$uniform(0,1)$	$uniform(.2,1.2)$	18	$\begin{pmatrix}0.66841\\0.59597\end{pmatrix}$
$uniform(0,1)$	$uniform(.1,1.1)$	60	$\begin{pmatrix}0.7263\\0.56576\end{pmatrix}$
$uniform(0,1)$	$uniform(.05,1.05)$	200	$\begin{pmatrix}0.75744\\0.52091\end{pmatrix}$

References

[1] C. B. Bell and K. A. Doksum, *Some new distribution-free statistics*, The Annals of Mathematical Statistics **36** (1965), no. 1, pp. 203–214 (English).