

We first make the following definitions:

$$X = [x_1, x_2, \dots, x_n]$$

$$Y = [y_1, y_2, \dots, y_m]$$

$$X = [X_T, X_V]$$

$$Y = [Y_T, Y_V]$$

That is X and Y are composed of observations, and each can be partitioned into training and validation sets.

$$\text{Define } f(X_v; X_t | h) \equiv \prod_{j \in X_v} \frac{1}{kh} \sum_{i \in X_t} K\left(\frac{j-i}{h}\right)$$

Where $k \equiv$ the number of elements in X_t

Recall our problem is we want to compute:

$$\text{BF} \equiv \frac{\int_0^\infty f([X_V, Y_V]; [X_T, Y_T] | h) f_1(h) dh}{\int_0^\infty f(Y_V; Y_T | h) f_3(h) dh \int_0^\infty f(X_V; X_T | h) f_2(h) dh}$$

and show this goes to 0 under some conditions and infinity in some other conditions.

We know a closed form for these integrals for a particular choice of f_1, f_2, f_3 and the Gaussian Kernel.

Thus we can say the BF is also equal to:

$$= \left(\frac{1}{(k_1 + k_2)\sqrt{2\pi}} \right)^{n_1 + n_2} \frac{B_1}{\sqrt{\pi}} \sum_{i_{n_1+n_2} \in [X_T, Y_T]} \sum_{i_{n_1+n_2-1} \in [X_T, Y_T]} \cdots \sum_{i_2 \in [X_T, Y_T]} \sum_{i_1 \in [X_T, Y_T]} \frac{\Gamma\left(\frac{n_1+n_2-1}{2}\right)}{(.5(2B_1^2 + \sum_{l=1}^{n_1+n_2} ([X_V, Y_V]_L - i_L)^2))^{\frac{(n_1+n_2)-1}{2}}}$$

divided by

$$\left(\frac{1}{(k_1)\sqrt{2\pi}} \right)^{n_1} \frac{B_2}{\sqrt{\pi}} \sum_{i_{n_1} \in [X_T]} \sum_{i_{n_1-1} \in [X_T]} \cdots \sum_{i_2 \in [X_T]} \sum_{i_1 \in [X_T]} \frac{\Gamma\left(\frac{n_1-1}{2}\right)}{(.5(2B_2^2 + \sum_{l=1}^{n_1} ([X_V]_L - i_L)^2))^{\frac{n_1-1}{2}}} *$$

$$\left(\frac{1}{(k_2)\sqrt{2\pi}} \right)^{n_2} \frac{B_3}{\sqrt{\pi}} \sum_{i_{n_2} \in [Y_T]} \sum_{i_{n_2-1} \in [Y_T]} \cdots \sum_{i_2 \in [Y_T]} \sum_{i_1 \in [Y_T]} \frac{\Gamma\left(\frac{n_2-1}{2}\right)}{(.5(2B_3^2 + \sum_{l=1}^{n_2} ([Y_V]_L - i_L)^2))^{\frac{n_2-1}{2}}}$$

Where

$k_1 \equiv$ the number of elements in X_T

$k_2 \equiv$ the number of elements in Y_T

$n_1 \equiv$ the number of elements in X_V

$n_2 \equiv$ the number of elements in Y_V

When we say $[Y_V]_L$, we mean Y_V is a set of elements, and we arbitrarily index them.

This can be simplified partially into:

$$\frac{k_1^{n_1} k_2^{n_2}}{(k_1 + k_2)^{n_1 + n_2}} \frac{B_1}{B_2 B_3 \sqrt{\pi}} * \frac{1}{\mathbf{B}\left(\frac{n_1-1}{2}, \frac{n_2-1}{2}\right)}$$

$$\sum_{i_{n_1+n_2} \in [X_T, Y_T]} \sum_{i_{n_1+n_2-1} \in [X_T, Y_T]} \cdots \sum_{i_2 \in [X_T, Y_T]} \sum_{i_1 \in [X_T, Y_T]} \frac{1}{(.5(2B_1^2 + \sum_{l=1}^{n_1+n_2} ([X_V, Y_V]_L - i_L)^2))^{\frac{(n_1+n_2)-1}{2}}} \text{ divided by}$$

$$\sum_{i_{n_1} \in [X_T]} \sum_{i_{n_1-1} \in [X_T]} \cdots \sum_{i_2 \in [X_T]} \sum_{i_1 \in [X_T]} \frac{1}{(.5(2B_2^2 + \sum_{l=1}^{n_1} ([X_V]_L - i_L)^2))^{\frac{n_1-1}{2}}} *$$

$$\sum_{i_{n_2} \in [Y_T]} \sum_{i_{n_2-1} \in [Y_T]} \cdots \sum_{i_2 \in [Y_T]} \sum_{i_1 \in [Y_T]} \frac{1}{(.5(2B_3^2 + \sum_{l=1}^{n_2} ([Y_V]_L - i_L)^2))^{\frac{n_2-1}{2}}}$$

B refers to the Beta function. This should just be a generalized Binomial coefficient.

We can view the sums with their k^n terms as averages instead of sums. Then the big sum with the constant can be interpreted as the average of the euclidean distance from the training set to the validation set over all possible choices of the training set if I drew them with replacement, with mild pertubation from the prior.

For now because of the assumptions of our test, $[X_V]_L$ is the same distribution as i_L .

Define $f \equiv 1/(B_1^2 + .5n * (x^2))$ where x is a random variable drawn from this distribution $[X_V]_L - i_L$

The taylor series for f(x) is: $\sum_{D=0}^\infty \frac{(.5n)^D x^{2D}}{B_1^{2D+1}}$ if $|x| < \frac{B_1}{.5n}$

if $|x| > \frac{B_1}{.5n}$ the appropriate taylor series is: $\sum_{D=1}^\infty \frac{(.5n)^{-D} x^{-2D}}{B_1^{2*(D-1)}}$ if $|x| < \frac{B_1}{.5n}$

Applying expectation to the first implies:

$$E[f(x)] = \sum_{D=0}^\infty \frac{(.5n)^D E[x^{2D}]}{B_1^{2D+2}}$$

The second is NOT dangerous. While the negative moments are normally not defined, since we are looking at a variable that is 0 for a large interval around 0, all the negative moments are infact GURANTEED to be defined. I will have to think about this a bit more later... This is good however, we no longer need to restrict the prior at all.

There's actually a similar story with the positive moments as well. I am under the impression all positive moments should exist because these are truncated moments, and not actual moments. Even the Cauchy distribution's moments should be finite, as this isn't integrating this from -infinity to infinity.

Define X and Y to be from the same distribution and independent then:

$$E[x^{2D}] = E[(X - Y)^{2D}] = \sum_{i=0}^{2D} \binom{2D}{i} E[X^{2D-i}] E[Y^i]$$

So...

$$\sum_{i_{n_1} \in [X_T]} \sum_{i_{n_1-1} \in [X_T]} \cdots \sum_{i_2 \in [X_T]} \sum_{i_1 \in [X_T]} \frac{1}{(.5(2B_2^2 + \sum_{l=1}^{n_1} ([X_V]_L - i_L)^2))^{\frac{n_1-1}{2}}} \text{ should be the same as:}$$

$$\sum_{i_{n_1} \in [X_T]} \sum_{i_{n_1-1} \in [X_T]} \cdots \sum_{i_2 \in [X_T]} \sum_{i_1 \in [X_T]} \left(\sum_{D=0}^{\infty} \frac{(.5n_1)^D \sum_{i=0}^{2D} \binom{2D}{i} E[X^{2D-i}] E[X^i]}{B_1^{2D+2}} \right)^{\frac{n_1-1}{2}}$$

Similarly:

$$\sum_{i_{n_2} \in [Y_T]} \sum_{i_{n_2-1} \in [Y_T]} \cdots \sum_{i_2 \in [Y_T]} \sum_{i_1 \in [Y_T]} \frac{1}{(.5(2B_3^2 + \sum_{l=1}^{n_2} ([Y_V]_L - i_L)^2))^{\frac{n_2-1}{2}}} \text{ should be the same as:}$$

$$\sum_{i_{n_2} \in [Y_T]} \sum_{i_{n_2-1} \in [Y_T]} \cdots \sum_{i_2 \in [Y_T]} \sum_{i_1 \in [Y_T]} \left(\sum_{D=0}^{\infty} \frac{(.5n_2)^D \sum_{i=0}^{2D} \binom{2D}{i} E[Y^{2D-i}] E[Y^i]}{B_2^{2D+2}} \right)^{\frac{n_2-1}{2}}$$

Finally:

$$\sum_{i_{n_1+n_2} \in [X_T, Y_T]} \sum_{i_{n_1+n_2-1} \in [X_T, Y_T]} \cdots \sum_{i_2 \in [X_T, Y_T]} \sum_{i_1 \in [X_T, Y_T]} \frac{1}{(.5(2B_1^2 + \sum_{l=1}^{n_1+n_2} ([X_V, Y_V]_L - i_L)^2))^{\frac{(n_1+n_2)-1}{2}}}$$

If distribution of X and Y are the same then:

$$\sum_{i_{n_1+n_2} \in [X_T, Y_T]} \sum_{i_{n_1+n_2-1} \in [X_T, Y_T]} \cdots \sum_{i_2 \in [X_T, Y_T]} \sum_{i_1 \in [X_T, Y_T]} \left(\sum_{D=0}^{\infty} \frac{((.5(n_1+n_2))^D \sum_{i=0}^{2D} \binom{2D}{i} E[Y^{2D-i}] E[X^i])}{B_3^{2D+2}} \right)^{\frac{n_1+n_2-1}{2}}$$

Earlier we said that we could treat these sums like means, then:

$$\frac{1}{(k_1+k_2)^{n_1+n_2}} \sum_{i_{n_1+n_2} \in [X_T, Y_T]} \sum_{i_{n_1+n_2-1} \in [X_T, Y_T]} \cdots \sum_{i_2 \in [X_T, Y_T]} \sum_{i_1 \in [X_T, Y_T]} \sum_{D=0}^{\infty} \frac{((.5(n_1+n_2))^D \sum_{i=0}^{2D} \binom{2D}{i} E[Y^{2D-i}] E[X^i])}{B_3^{2D}} \text{ should be the same as}$$

$$\left(\sum_{D=0}^{\infty} \frac{((.5(n_1+n_2))^D \sum_{i=0}^{2D} \binom{2D}{i} E[Y^{2D-i}] E[X^i])}{B_3^{2D+2}} \right)^{\frac{n_1+n_2-1}{2}}$$

Alternatively this is true because the term inside our sum no longer depends on i_l so it is the same as adding a constant many times and then dividing by the number of times we summed.

Similarly:

$$\frac{1}{(k_2)^{n_2}} \sum_{i_{n_2} \in [Y_T]} \sum_{i_{n_2-1} \in [Y_T]} \cdots \sum_{i_2 \in [Y_T]} \sum_{i_1 \in [Y_T]} \left(\sum_{D=0}^{\infty} \frac{(.5n_2)^D \sum_{i=0}^{2D} \binom{2D}{i} E[Y^{2D-i}] E[Y^i]}{B_2^{2D+2}} \right)^{\frac{n_2-1}{2}}$$

is the same as:

$$\left(\sum_{D=0}^{\infty} \frac{(.5n_2)^D \sum_{i=0}^{2D} \binom{2D}{i} E[Y^{2D-i}] E[Y^i]}{B_2^{2D+2}} \right)^{\frac{n_2-1}{2}}$$

and

$$\frac{1}{(k_1)^{n_1}} \sum_{i_{n_1} \in [X_T]} \sum_{i_{n_1-1} \in [X_T]} \cdots \sum_{i_2 \in [X_T]} \sum_{i_1 \in [X_T]} \left(\sum_{D=0}^{\infty} \frac{(.5n_1)^D \sum_{i=0}^{2D} \binom{2D}{i} E[X^{2D-i}] E[X^i]}{B_1^{2D+2}} \right)^{\frac{n_1-1}{2}}$$

is around:

$$\left(\sum_{D=0}^{\infty} \frac{(.5n_1)^D \sum_{i=0}^{2D} \binom{2D}{i} E[X^{2D-i}] E[X^i]}{B_1^{2D+2}} \right)^{\frac{n_1-1}{2}}$$

Our BF under the Null then is roughly:

$$\frac{B_1}{B_2 B_3 \sqrt{\pi}} * \frac{1}{B(\frac{n_1-1}{2}, \frac{n_2-1}{2})} *$$

$$\sum_{D=0}^{\infty} \frac{((.5(n_1+n_2))^D \sum_{i=0}^{2D} \binom{2D}{i} E[Y^{2D-i}] E[X^i])}{B_3^{2D+2}} \text{ divided by}$$

$$\left(\sum_{D=0}^{\infty} \frac{(.5n_2)^D \sum_{i=0}^{2D} \binom{2D}{i} E[Y^{2D-i}] E[Y^i]}{B_2^{2D+2}} \right) * \left(\sum_{D=0}^{\infty} \frac{(.5n_1)^D \sum_{i=0}^{2D} \binom{2D}{i} E[X^{2D-i}] E[X^i]}{B_1^{2D+2}} \right)$$

The B_i 's are from the prior, and is the IQR, if the distributions are the same we expect them to be similar, so... the BF under null should be

$$\frac{1}{B_1 \sqrt{\pi}} * \frac{1}{\mathbf{B}(\frac{n_1-1}{2}, \frac{n_2-1}{2})} * (\sum_{D=0}^{\infty} \frac{((.5(n_1+n_2))^D \sum_{i=0}^{2D} \binom{2D}{i} E[X^{2D-i}] E[X^i])}{B_1^{2D+2}})^{\frac{n_1+n_2-1}{2}} \text{ divided by } (\sum_{D=0}^{\infty} \frac{(.5n_2)^D \sum_{i=0}^{2D} \binom{2D}{i} E[X^{2D-i}] E[X^i]}{B_1^{2D+2}})^{\frac{n_2-1}{2}} * (\sum_{D=0}^{\infty} \frac{(.5n_1)^D \sum_{i=0}^{2D} \binom{2D}{i} E[X^{2D-i}] E[X^i]}{B_1^{2D+2}})^{\frac{n_1-1}{2}}$$

For simplicity we show first what happens when $n_1 = n_2$

$$\frac{1}{B_1 \sqrt{\pi}} * \frac{1}{\mathbf{B}(\frac{n_1-1}{2}, \frac{n_1-1}{2})} * (\sum_{D=0}^{\infty} \frac{(((n_1))^D \sum_{i=0}^{2D} \binom{2D}{i} E[X^{2D-i}] E[X^i])}{B_1^{2D+2}})^{n_1-.5} \text{ divided by } (\sum_{D=0}^{\infty} \frac{(.5n_1)^D \sum_{i=0}^{2D} \binom{2D}{i} E[X^{2D-i}] E[X^i]}{B_1^{2D+2}})^{n_1-1}$$

$$\text{An obvious inequality is: } (\sum_{D=0}^{\infty} \frac{(((n_1))^D \sum_{i=0}^{2D} \binom{2D}{i} E[X^{2D-i}] E[X^i])}{B_1^{2D+2}})^{n_1-.5} > (\sum_{D=0}^{\infty} \frac{(.5n_1)^D \sum_{i=0}^{2D} \binom{2D}{i} E[X^{2D-i}] E[X^i]}{B_1^{2D+2}})^{n_1-.5}$$

So this part of the BF is lowerbounded by:

$$\frac{1}{B_1 \sqrt{\pi}} * \frac{1}{\mathbf{B}(\frac{n_1-1}{2}, \frac{n_1-1}{2})} * (\sum_{D=0}^{\infty} \frac{(((n_1))^D \sum_{i=0}^{2D} \binom{2D}{i} E[X^{2D-i}] E[X^i])}{B_1^{2D+2}})^{.5}$$

Recall the following definitions...

$f \equiv 1/(B_1^2 + .5n * (x^2))$ where x is a random variable drawn from this distribution $[X_V]_L - i_L$

$$E[f(x)] = \sum_{D=0}^{\infty} \frac{(.5n)^D \sum_{i=0}^{2D} \binom{2D}{i} E[X^{2D-i}] E[X^i]}{B_1^{2D+2}}$$

Then the sum is lowerbounded by $E[f(x)]$ where $f \equiv 1/(B_1^2 + n_1 * (x^2))$ where x is a random variable drawn from this distribution $[X_V]_L - i_L$

Whatever the case, for the lower bound, the Beta term rises exponentially with n , and the sum also increases with n , but I don't know the exact amount. As a result, the lowerbound for this BF should go to infinity at least exponentially. Which implies the BF itself should also go to infinity at least exponentially under the null hypothesis.

This method works if $|x| < B_1$.

Under the alternate we would have to split up the sum into parts where we subtract over the same and different parts. That is introduce indicators.

Nothing changes with the denominator in the BF, the numerator should now have a very interesting expression. Regardless, if two distributions are different it'll show in their moments! So things will behave very differently compared to the null situation.

$$\sum_{i_{n_1+n_2} \in [X_T, Y_T]} \sum_{i_{n_1+n_2-1} \in [X_T, Y_T]} \cdots \sum_{i_2 \in [X_T, Y_T]} \sum_{i_1 \in [X_T, Y_T]} \frac{1}{(.5(2B_1^2 + \sum_{L=1}^{n_1+n_2} ([X_V, Y_V]_L * I_{([X_V, Y_V]_L \in X_V) - i_L}^2)))^{\frac{(n_1+n_2)-1}{2}}}$$