(2/9) Yesterday I decided to look at the auc's of each coefficient estimate separately, analogous to the alternative hypothesis project. that is, obtain a second order expansion of $\theta(\hat{F}, \hat{G}, \hat{\beta}) - \theta(F, G, \beta)$ and similarly for $\hat{\gamma}$, then take the difference. the constant and linear terms will cancel under the null. yesterday i verified the 2nd order taylor expansion due to $\hat{\beta} - \beta$, and since then the function expansion due to $\hat{F} - F$, and G. i reviewed the Lee material on degenerate U-stats, since that is what the residual term in $\theta(\hat{F}, \hat{G}, \hat{\beta}) -$ const term $-$ hajek/linear term, but couldn't figure out how to get the expected chi squared mixture.

Earlier today I noticed $n \int \psi(x, y) d(\hat{F} - F) d(\hat{G} - G) \to \int \psi(x, y) dB_1 dB_2$ where B_i are independent brownian bridges. why is this the same as a combination of chi squares?

I then looked at the Heller paper to see what htey did under the null, and realized since teh const terms are the smae under the null, i can just look at the expansion of $\theta(\hat{F}, \hat{G}, \hat{\beta}) - \theta(\hat{F}, \hat{G}, \hat{\gamma})$. So in this situation the differnce of AUCs is easier than computing the auc expansions separately then taking the difference.

so an analysis v similar to heller's might work, at least if the coefficient estimation model is something like well specified logistic or lda. the advantage heller cites of mrc is that the first order term vanishes, but the same holds here too. (2/11) Realized cant directly copy heller approach. the key fact about mrc is that the auc vanishes at the estimated coefficient, not the plim coefficient, as with well specified logistic or lda. that also explains why the proof expands around thetahat rather than that. However, from 1e it seems that the 2nd order hoeffding term $\theta(\delta F, \delta G, \beta)$ is $o(1/n)$ when β is random, $\beta + z_i/\sqrt{n}$. When the normalizer is smaller or larger than $\sqrt{n}$ the rate drops back to $O(1/n)$...in particular when β is nonrandom, no noise z_i is added.