STAT C206A / MATH C223A : Stein's method and applications 1

# Lecture 33

*Lecture date: Nov 14, 2007* *Scribe: Maximilian Kasy*

## 1 Zero Bias Coupling

Consider a random variable $W$ with $EW = 0$ and $EW^2 = \sigma^2 < \infty$. Define

$$\rho(x) = \frac{1}{\sigma^2} E(W 1_{(W \geq x)}).$$

**Lemma 1.** *$\rho$ is a probability density*

Proof:

- $\rho(x) \geq 0$ for all $x$. For $x \geq 0$ this is obvious, for $x < 0$ note $E(W 1_{(W \geq x)}) = -E(W 1_{(W<x)}) \geq 0$
- $\int \rho(x) dx = 1$:

  $$\int \rho(x) dx = \int_{x<0} -E(W 1_{(W<x)}) dx + \int_{x \geq 0} E(W 1_{(W \geq x)}) dx$$

  Let $\mu$ denote the law of $W$. Then

  $$\int_{x \geq 0} E(W 1_{(W \geq x)}) dx = \int_{x \geq 0} \int_{y \geq x} y d\mu(y) dx = \int_{y \geq 0} \int_{x \leq y} y \; dx \; d\mu(y) = \int_{y \geq 0} y^2 d\mu(y)$$

  and similarly for the second term.□

The distribution corresponding to this density is called the "zero bias transform". If $W^*$ is a random variable following the zero-bias transform of the law of $W$, then for all absolutely continuous $\phi$ we have

$$EW\phi(W) = \sigma^2 E\phi'(W^*)$$

(this is immediate from integration by parts). If $W \sim N(0,1)$ then

$$\rho(x) = \int_{y \geq x} \frac{y e^{-y^2/2}}{\sqrt{2\pi}} dy = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$

so that $W$ $W^*$, i.e. the standard normal distribution is a fixed point of the zero bias transform.

Example: $W \sim \pm 1$. Then $W^* \sim Uni[-1, 1]$.

**Theorem 1.** *If $EW = 1$, $EW^2 = 1$, then* $Wass(W, Z) \leq 2Wass(W, W^*)$

Proof: Take any 1-Lipschitz $h$. Find $\phi$ such that $\phi'(x) - x\phi(x) = h(x) - Eh(Z)$. By earlier results we know $\|\phi''\|_\infty \leq 2\|h'\|_\infty \leq 2$. Now suppose $W$ and $W^*$ live on the same space. Then

$$|E(h(W)) - E(h(Z))| = |E(\phi'(W) - W\phi(W))| = |E(\phi'(W) - \phi'(W^*))| \leq 2E|W^* - W|$$

SInce this is true for any coupling, any $h$ Lipschitz we have

$$Wass(W, Z) \leq 2 \inf_{W, W^*} E|W^* - W| = 2Wass(W, W^*)$$

Example: suppose $X_1, \ldots, X_n$ i.i.d. mean 0 variance 1, $W = \frac{\sum X_i}{\sqrt{n}}$. For each $i$ let $X_i$ be independent of everything else. Let $I \sim Unif\{1, \ldots, n\}$ independent of everything else. Define

$$W^* := \frac{1}{\sqrt{n}} \left[ \sum_{i \neq I} X_i + X_I^* \right]$$

We claim this is a zero bias transform:

$$EW\phi(W) = \frac{1}{\sqrt{n}} \sum EX_i\phi(W) =$$

$$\frac{1}{\sqrt{n}} \sum E\left( X_i\phi\left( \frac{1}{\sqrt{n}} \left[ \sum_{j \neq i} X_j + X_i \right] \right) \right)$$

$$\frac{1}{\sqrt{n}} \sum E\left( \phi'\left( \frac{1}{\sqrt{n}} \left[ \sum_{j \neq i} X_j + X_i^* \right] \right) \frac{1}{\sqrt{n}} \right)$$

$$= E(\phi'(W^*))$$

Thus

$$Wass(W, Z) \leq 2E|W^* - W| = 2E|\frac{1}{\sqrt{n}}(X_I - X_I^*)| = 2E|\frac{1}{\sqrt{n}}(X_1 - X_1^*)| \leq \ldots$$

**Theorem 2.** *(Goldstein-Reinert) Suppose $Y, Y'$ are an exchangeable pair, $EY = 0$, $EY^2 = \sigma^2$ and $E(Y'|Y) = (1 - \lambda)Y$. Let $\nu$ denote the joint distribution of $(Y, Y')$. Let*

$$d\mu(y, y') = \frac{(y - y')^2}{E(Y - Y')^2} d\nu(y, y').$$

*Suppose $(\hat{Y}, \hat{Y}') \sim \mu$. Let $U \sim Uni[0, 1]$ independent of all else. $Y^* = U\hat{Y} + (1 - U)\hat{Y}'$. Then $Y^*$ is a zero bias transform of $Y$.*

Note that $Ef'(Ua+(1-U)b) = \frac{f(b)-f(a)}{b-a}$, hence

$$\sigma^2 E(f'(Y^*)) = \sigma^2 E\left(\frac{f(\hat{Y}) - f(\hat{Y}')}{\hat{Y} - \hat{Y}'}\right)$$

$$= \frac{\sigma^2}{E(Y-Y')^2} E[(Y-Y')(f(Y)-f(Y'))] = \frac{2\lambda\sigma^2 E[Yf(Y)]}{E(Y-Y')^2} = EYf(Y)$$

Exercise: Try to get Hoeffding CLT using this method.

Exercise: Suppose $EW\phi(W) = ET\phi'(W)$.Then $E(T|W)$ is the density of $Law(W^*)$ w.r.t. $Law(W)$ evaluated at $W$. We have Tusnády's lemma based on concentration of $T$. On the other hand, if $W, W^*$ can be constructed to be close to each other then we can construct $W, Z$ such that $E|W-Z|$ is small. Question: If we know tail bounds on $|W-W^*|$ can we construct $(W,Z)$ with fast decaying tails for $W-Z$?