1 Bin packing.

Bin packing. This is a classical application of McDiarmid's inequality. Let $X_1, \ldots, X_n$ be i.i.d. random variables taking values in $[0, 1]$. Each X_i is the size of a package to be shipped. The packages are shipped in bins of size 1, so each bin can hold any set of packages whose sizes sum to at most 1. Let $B_n = f(X_1, \ldots, X_n)$ be the minimal number of bins needed to ship the packages with sizes $X_1, \ldots, X_n$. Computing B_n is a hard combinatorial optimization problem; however, we can say something about its mean and tail behavior.

1.1

Let μ be the common mean of the X_i's. Show that $\mathbb{E}[B_n] \geq n\mu$.

1.2

Prove that for any $\epsilon > 0$,

$$\mathbb{P}\left(n^{-1}B_n \leq \mu - \epsilon\right) \leq \exp\left(-2n\epsilon^2\right).$$

2 Orthogonal design lasso bits

Let $y_i = x_i^\top \beta^* + \eta_i$ for $i = 1, \ldots, n$. Assume that the design matrix $\mathbf{X} = (x_1, \ldots, x_n)^\top$ is not random. Also assume that $\log \mathbb{E}\left[e^{\lambda \eta_i}\right] \leq \sigma^2\lambda^2/2$. Let

$$\widehat{\beta} = \operatorname*{argmin}_{\beta} \frac{1}{2n}\|Y - \mathbf{X}\beta\| + \lambda\|\beta\|_1.$$

2.1

Derive the so-called "basic inequality":

$$\|\mathbf{X}\beta^* - \mathbf{X}\widehat{\beta}\| + \lambda\|\beta\|_1 \leq \frac{1}{n}\eta^\top \mathbf{X}(\widehat{\beta} - \beta^*) + \lambda\|\beta^*\|_1$$

2.2

Let X_j be the j^{th} column of $\mathbf{X}$. Find a constant M such that $\mathbb{P}(\mathcal{G}) \geq 1 - 2e^{-\delta}$ for

$$\mathcal{G} = \left\{\max_j |\eta^\top X_j|/n < M\right\}.$$

3 Empirical Rademacher Complexity

Prove the following result. Let H be a set of hypotheses bounded by $M > 0$ and consider an i.i.d. sample of size n. For any $h \in H$, with probability at least $1 - \delta$

$$R_n(h) \leq \widehat{R}_n(h) + \widehat{\mathcal{R}}_n(H) + 3M\sqrt{\frac{2\log(2/\delta)}{n}}.$$

You may use the result from class that

$$R_n(h) \leq \widehat{R}_n(h) + \mathcal{R}_n(H) + M\sqrt{\frac{2\log(1/\delta)}{n}}.$$