

Sidework #1
RELEASE DATE: 10/04/2011
DUE DATE: NONE

1.1 Hoeffding's Inequality

The proof that you will write below contains all the essential steps, but are not as rigorously written as the usual math texts.

- (1) (Markov's Inequality) Prove that for any non-negative random variable ℓ and any positive constant a ,

$$P(\ell \geq a) \leq \frac{E(\ell)}{a}.$$

- (2) (Moment-Generating Function—Laplace Transform) Prove that for any finite random variable λ , any positive constant α , and any positive parameter s ,

$$\begin{aligned} P(\lambda \geq \alpha) &\leq e^{-s\alpha} E(e^{s\lambda}); \\ P(\lambda \leq \alpha) &\leq e^{s\alpha} E(e^{-s\lambda}). \end{aligned}$$

- (3) (Independence Decomposition) Let z_1, z_2, \dots, z_N be i.i.d. random variables and let $z = \frac{1}{N} \sum_{n=1}^N z_n$. For any positive constant α and any positive parameter s , prove that

$$\begin{aligned} P(z \geq \alpha) &\leq (e^{-s\alpha} E(e^{sz_1}))^N; \\ P(z \leq \alpha) &\leq (e^{s\alpha} E(e^{-sz_1}))^N. \end{aligned}$$

- (4) (Bound Tightening) Let z_1 be a binary random variable with $P(z_1 = 0) = 1 - \theta$ and $P(z_1 = 1) = \theta$. Let

$$F(s) = e^{-s\alpha} E(e^{sz_1})$$

For any given α with $\theta < \alpha < 1$, prove that $F(s)$ is minimized on

$$s^* = \ln \frac{\alpha \cdot (1 - \theta)}{(1 - \alpha) \cdot \theta},$$

where s^* is positive.

- (5) (Chernoff Bound) Use the fact that

$$P(z \geq \alpha) \leq (e^{-s^*\alpha} E(e^{s^*z_1}))^N$$

to prove

$$P(z \geq \alpha) \leq e^{-ND(\alpha||\theta)}$$

for $\theta < \alpha < 1$. Here $D(\alpha||\theta) = \alpha \log \frac{\alpha}{\theta} + (1 - \alpha) \log \frac{1 - \alpha}{1 - \theta}$ is the KL-divergence between α and θ .

- (6) (One-sided Hoeffding Inequality) Let $\alpha = \theta + \epsilon$ with $\epsilon > 0$, prove that

$$P(z - \theta \geq \epsilon) \leq e^{-2N\epsilon^2}$$

by showing that $D(\theta + \epsilon||\theta) \geq 2\epsilon^2$ for all $\epsilon > 0$.

- (7) (Two-sided Hoeffding Inequality) Prove that

$$P(|z - \theta| \geq \epsilon) \leq 2e^{-2N\epsilon^2}.$$