Statistics 164 Homework 4

1. Let $X \sim \text{Bin}(n, p)$. Find the MGF of X. Show that if $p \geq 1/2$ and $\epsilon \leq 1 - p$, then

$$P(X - np \;\geq\; n\epsilon) \;\leq\; \exp\left\{\frac{-n\epsilon^2}{2p(1-p)}\right\}$$

Hint: Use the Chernoff bound, with an optimal choice of the free parameter, and the inequality

$$x \log \frac{x}{p} \;+\; (1-x) \log \frac{1-x}{1-p} \;\geq\; \frac{(x-p)^2}{2p(1-p)}$$

Optional: verify that the inequality holds for the appropriate range of x and $p \geq 1/2$.

2. Let $h(u)$ be the function appearing in the version of Bennett's inequality given in class. Show, as claimed, that $h(u) \geq u^2(2 + 2u/3)^{-1}$.

3. Let $h(u) = -u \log u - (1-u) \log(1-u)$ be the binary entropy function.

   a. Show that $h(u)$ is continuous and concave on $[0, 1]$ with a maximum at $1/2$.

   b. Show that $h(u)$ is symmetric about $1/2$.

   c. Use the bound on $\binom{n}{m}$ obtained in the last HW to show that for fixed $\alpha \in (0, 1)$

$$\lim_{n\to\infty} \frac{1}{n} \log_2 \binom{n}{\alpha n} \;=\; h(\alpha)$$

4. Let X, Y, Z be jointly distributed discrete random variables. Show that $H(X, Y) = H(X) + H(Y|X)$ and that $H(X|Y) \geq H(X|Y, Z)$. For the second inequality, you can condition on the value of Y and use the fact that $H(X) \geq H(X|Y)$.

5. Show that $H(X_1, \ldots, X_n) \leq H(X_1) + \cdots + H(X_n)$.

6. Let (X, Y) be jointly distributed r.v. with $EY^2 < \infty$. The conditional variance of Y given X is defined by $\text{Var}(Y|X) = E[(Y - E(Y|X))^2|X]$.

   (a) Show that $\text{Var}(Y|X) = E(Y^2|X) - (E(Y|X))^2$.

   (b) Show that $\text{Var}[E(Y|X)] = E[E(Y|X)^2] - (E(Y))^2$.

   (c) Use (a), (b) to show that $\text{Var}(Y) = E[\text{Var}(Y|X)] + \text{Var}[E(Y|X)]$.

7. [Steele] Recall that the Euclidean norm of a vector $a = (a_1, \ldots, a_m) \in \mathbb{R}^m$ is given by $||a|| = \sqrt{a_1^2 + \cdots + a_m^2}$. The Cauchy-Schwartz (CS) inequality states that for $a, b \in \mathbb{R}^m$,

$$a^T b \;=\; a_1 b_1 + \cdots + a_m b_m \;\leq\; ||a||\,||b||.$$

a. Prove the CS inequality starting, from the basic fact that

$$0 \leq \left(\frac{a_j}{||a||} - \frac{b_j}{||b||} \right)^2$$

for each $j = 1, \ldots, m$.

b. Use the CS inequality to establish that

$$\sum_{i=1}^{n} a_i \;\leq\; n^{1/2} \, (\sum_{i=1}^{n} a_i^2)^{1/2}$$

and that

$$\sum_{i=1}^{n} a_i \;\leq\; (\sum_{i=1}^{n} |a_i|^{2/3})^{1/2} \, (\sum_{i=1}^{n} |a_i|^{4/3})^{1/2}$$

c. Show that

$$(\sum_{i=1}^{n} a_i b_i c_i)^2 \;\leq\; (\sum_{i=1}^{n} a_i^2) \, (\sum_{i=1}^{n} b_i^2) \, (\sum_{i=1}^{n} c_i^2)$$

(Hint: Bound the numbers c_i by their maximum and apply CS.) Apply the bound with $a_i = b_i = c_i = 1$. What does this tell you about how "good" the bound is?

d. Show that for $x, y, z > 0$ we have the inequality

$$x + y + z \;\leq\; 2 \left(\frac{x^2}{y+z} + \frac{y^2}{x+z} + \frac{z^2}{x+y} \right)$$

8. Let X, Y, Z be discrete random variables. Show that $E[\,E(Z \,|\, Y, X) \,|\, Y] = E(Z \,|\, Y)$.

9. Let f be a convex function on the interval $[a, b]$ and let $a < x < b$. Show that

$$\frac{f(x) - f(a)}{x - a} \;\leq\; \frac{f(b) - f(a)}{b - a} \;\leq\; \frac{f(b) - f(x)}{b - x}.$$

(Hint: express x as a convex combination of a and b and then apply the definition of convexity.) Draw a picture illustrating this result, and give a brief explanation.

10. Show that

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} \;=\; 1.$$

Hint: Rewrite each term in the sum as a difference $(a_k - a_{k+1})$, and then cancel like terms.