PROBLEMS - SET 5

Problem 1. Let (X_n) be a i.i.d. sequence of $Bin(1,p)$. Prove the following Chernoff bounds:

$$\mathrm{P}(\overline{X}_n \geq (1+\delta)p) \leq e^{-np\frac{\delta^2}{2+\delta}},$$

and

$$\mathrm{P}(\overline{X}_n \leq (1-\delta)p) \leq e^{-np\frac{\delta^2}{2}}.$$

Solution 1. See Lecture Notes, pag. 56

Problem 2. Let (X_n) be a i.i.d. sequence of $N(0,1)$.

(a) Prove the following Chernoff bound:

$$\mathrm{P}(\overline{X}_n \geq \varepsilon) \leq e^{-n\frac{\varepsilon^2}{2}}$$

(b) Obtain a sharper upper tail estimate, proving first the inequality

$$\mathrm{P}(X_1 \geq x) \leq \frac{1}{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

and deriving then that

$$\mathrm{P}(\overline{X}_n \geq \varepsilon) \leq \frac{1}{\varepsilon\sqrt{n}}\frac{1}{\sqrt{2\pi}}e^{-n\frac{\varepsilon^2}{2}}$$

(c) Prove the further inequality

$$\mathrm{P}(X_1 \geq x) > \left(\frac{1}{x} - \frac{1}{x^3}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

Solution 2. (a) See Lecture 17

(b) See Lecture 17

(c) Let us prove that

$$\mathrm{P}(X_1 \geq x) > \left(\frac{1}{x} - \frac{1}{x^3}\right)\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2).$$

We have that $\Phi'(t) = -t\Phi(t)$ iff $\Phi(t) = -\frac{1}{t}\Phi'(t)$, and we get

$$\mathrm{P}(X_1 \geq x) = \int_x^{+\infty}\Phi(t)dt = -\int_x^{+\infty}\frac{1}{t}\Phi'(t)dt.$$

Solving the integral by parts

$$-\int_x^{+\infty}\frac{1}{t}\Phi'(t)dt = -\left[\frac{1}{t}\Phi(t)\right]_x^{+\infty} + \int_x^{+\infty}\Phi(t)\left(-\frac{1}{t^2}\right)dt = \frac{1}{x}\Phi(x) - \int_x^{+\infty}\frac{\Phi(t)}{t^2}dt.$$

Then

$$\int_x^{+\infty} \frac{\Phi(t)}{t^2} dt = \int_x^{+\infty} \frac{t\Phi(t)}{t^3} dt < \frac{1}{x^3} \int_x^{+\infty} t\Phi(t) dt = \frac{1}{x^3} \Phi(x).$$

and

$$\mathrm{P}(X_1 \geq x) > \frac{1}{x}\Phi(x) - \frac{1}{x^3}\Phi(x) = \left(\frac{1}{x} - \frac{1}{x^3}\right) \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}x^2).$$

Problem 3. Let (X_n) be a i.i.d. sequence of $\mathrm{Pois}(\lambda)$. Prove the following Chernoff bounds:

$$\mathrm{P}(\overline{X}_n \geq \lambda(1+\varepsilon)) \leq e^{-n\lambda a(\varepsilon)},$$

and, for $0 < \varepsilon < 1$

$$\mathrm{P}(\overline{X}_n \leq \lambda(1-\varepsilon)) \leq e^{-n\lambda b(\varepsilon)},$$

where

$$a(\varepsilon) = \frac{\varepsilon^2}{2+\varepsilon}$$

and

$$b(\varepsilon) = \frac{\varepsilon^2}{2}.$$

Hint: prove (and use) the following inequalities

$$(1+\varepsilon)\log(1+\varepsilon) - \varepsilon := f(\varepsilon) \geq a(\varepsilon) = \frac{\varepsilon^2}{2+\varepsilon}, \ \ \forall \varepsilon > 0, \tag{0.1}$$

$$\varepsilon + (1-\varepsilon)\log(1-\varepsilon) := k(\varepsilon) \geq b(\varepsilon) = \frac{\varepsilon^2}{2} \ \ \forall\, 0 < \varepsilon < 1. \tag{0.2}$$

Note that if F, G are C^2 functions in $[0,a)$, $F(0) = G(0)$, $F'(0) = G'(0)$ and $F''(x) \geq G''(x)$ for all $x \in [0,a)$, then $F(x) \geq G(x)$ for $x \in [0,a)$.

Solution 3. We first compute

$$m(t) = \mathrm{E}\left[e^{tX_1}\right] = \sum_{n=0}^{+\infty} e^{-\lambda} e^{tn} \frac{\lambda^n}{n!} = e^{\lambda(e^t-1)}.$$

So let

$$g(t) = t\lambda(1+\varepsilon) - \log m(t) = t\lambda(1+\varepsilon) - \lambda(e^t - 1).$$

The maximum of this function is attained at $t^* = \log(1+\varepsilon)$, and is

$$g(t^*) = \lambda\left[(1+\varepsilon)\log(1+\varepsilon) - \varepsilon\right].$$

Therefore we have that

$$\mathrm{P}(\overline{X}_n \geq \lambda(1+\varepsilon)) \leq e^{-ng(t^*)} = e^{-n\lambda\left[(1+\varepsilon)\log(1+\varepsilon)-\varepsilon\right]}.$$

So we are left to prove (0.1). Let $f(\varepsilon) = [(1+\varepsilon)\log(1+\varepsilon) - \varepsilon]$ and observe that

$$f'(\varepsilon) = \log(1+\varepsilon) \qquad f''(\varepsilon) = \frac{1}{1+\varepsilon}.$$

Note that

$$a'(\varepsilon) = \frac{2\varepsilon(2+\varepsilon) - \varepsilon^2}{(2+\varepsilon)^2} \qquad a''(\varepsilon) = \frac{8}{(2+\varepsilon)^3}.$$

We observe thatt $(2+\varepsilon)^3 - 8(1+\varepsilon) = 4\varepsilon + 6\varepsilon^2 + \varepsilon^3 = \varepsilon(4+6\varepsilon+\varepsilon^2) > 0$ if $\varepsilon > 0$. So $f''(\varepsilon) > a''(\varepsilon)$ for $\varepsilon > 0$.

Note that $f(0) = a(0) = 0$, $f'(0) = g'(0) = 0$ and $f''(\varepsilon) \geq g''(\varepsilon)$ for all $\varepsilon > 0$, since $(2+\varepsilon)^3 \geq 8(1+\varepsilon)$ for all $\varepsilon > 0$. Therefore $f(\varepsilon) \geq a(\varepsilon)$.

Similarly, for the lower tail, we define

$$h(t) = -t\lambda(1-\varepsilon) - \log m(-t) = -t\lambda(1-\varepsilon) - \lambda(e^{-t} - 1),$$

which is maximized at $t^* = -\log(1-\varepsilon)$, giving

$$h(t^*) = \lambda[\varepsilon + (1-\varepsilon)\log(1-\varepsilon)].$$

Then

$$\mathrm{P}(\overline{X}_n \leq \lambda(1-\varepsilon)) \leq e^{-nh(t^*)} = e^{-n\lambda[\varepsilon+(1-\varepsilon)\log(1-\varepsilon)]}.$$

So we are left to prove (0.2). Observe that

$$k'(\varepsilon) = -\log(1-\varepsilon), \;\; k''(\varepsilon) = \frac{1}{1-\varepsilon} \qquad b'(\varepsilon) = \varepsilon, \;\; b''(\varepsilon) = 1.$$

So $k(0) = b(0) = 0$, $k'(0) = g'(0) = 0$ and $k''(\varepsilon) > 1 = g''(\varepsilon)$ for all $\varepsilon \in [0,1)$. This implies the conclusion.

Problem 4. Let (X_n) be a i.i.d. sequence with distribution Unif$(-1,1)$

(a) Show that $m_{X_n}(t) = \frac{\sinh(t)}{t}$.

(b) Show that there exists $k > 0$ such that for all $0 \leq t \leq k$

$$\frac{\sinh(t)}{t} \leq 1 + \frac{t^2}{2}$$

(numerically $k \simeq 4.75$.)

(c) Prove the Chernoff bound: for all $0 < \varepsilon \leq k$

$$\mathrm{P}(\overline{X}_n \geq \varepsilon) \leq e^{-n\frac{\varepsilon^2}{2}}.$$

(d) Without any further computation, explain why the bound for the lower tail

$$\mathrm{P}(\overline{X}_n \leq -\varepsilon) \leq e^{-n\frac{\varepsilon^2}{2}}.$$

also holds for all $0 < \varepsilon \leq k$.

Solution 4. (a)

$$m(t) = m_{X_n}(t) = \frac{1}{2}\int_{-1}^{1} e^{tx}dx = \frac{e^t - e^{-t}}{2t} = \frac{\sinh(t)}{t}.$$

(b) Taylor expansion of $\sinh(t)$ at $t = 0$ gives

$$\sinh(t) = \sum_{k=0}^{+\infty} \frac{t^{2k+1}}{(2k+1)!} = t + \frac{t^3}{6} + \frac{t^5}{5!} + \ldots.$$

Note that

$$\sinh(t) - t - \frac{t^3}{2} = -\frac{t^3}{3} + \sum_{k=2}^{+\infty} \frac{t^{2k+1}}{(2k+1)!} = -\frac{t^3}{3} + t^5 \sum_{k=0}^{+\infty} \frac{t^{2k}}{(2k+5)!}.$$

Note that $\sum_{k=0}^{+\infty} \frac{t^{2k}}{(2k+5)!} < +\infty$ for every $t > 0$. For $t > 0$ sufficiently small it is easy to check that $-\frac{t^3}{3} + t^5 \sum_{k=0}^{+\infty} \frac{t^{2k}}{(2k+5)!} < 0$, and so the conclusion follows. To estimate roughly $k > 0$ such that

$$\frac{\sinh(t)}{t} \leq 1 + \frac{t^2}{2} \qquad \text{for all } t \in [0, k]$$

one could proceed as follows. Define $f(t) = \sinh t - t - t^3/2$ and note that $f(0) = 0$, $f'(t) = \cosh t - 1 - 3t^2/2$, $f'(0) = 0$, $f''(t) = \sinh t - 3t$, $f''(0) = 0$. So, for $t \leq 3$, $f''(t) < 0$ and then $f'(t) < 0$. Now observe that $f'(4) = \cosh 4 - 25 > 0$. So, there exists $c \in (3, 4)$ such that $f'(c) = 0$ and $f'(t) < 0$ for $t \in (0, c)$, and $f'(t) > 0$ for $t > c$. In particular that $f(t) < 0$ for $t \in (0, c)$ and that there exists $k > c$ such that $f(k) = 0$. Compute $f(4) = \sinh 4 - 4 - 32 < 0$ and $f(5) = \sinh 5 - 5 - 125/2 > 0$. So $k \in (4, 5)$.

(c) For $t \leq k$

$$g(t) := t\varepsilon - \log m(t) = t\varepsilon - \log \frac{\sinh(t)}{t} \geq t\varepsilon - \log\left(1 + \frac{t^2}{2}\right) \geq t\varepsilon - \frac{t^2}{2},$$

where we have used the inequality $\log(1 + x) \leq x$. The function $t\varepsilon - \frac{t^2}{2}$ is maximized at $t^* = \varepsilon$. So for $\sigma \leq k$

$$g(t^*) \geq t^*\varepsilon - \frac{(t^*)^2}{2} = \frac{\varepsilon^2}{2},$$

and the conclusion follows from the general Chernoff bounds.

(d) Note that $h(t) = t\varepsilon - \log m(-t) = t\varepsilon - \log \frac{\sinh(-t)}{-t} = g(t)$.

Problem 5. Let (X_n) be a i.i.d. sequence such that

$$\varphi_{X_n}(u) = \frac{1}{1+u^2}.$$

For every $\varepsilon > 0$ find $a(\varepsilon) > 0$ (no "nice" form is necessary) such that

$$P(\overline{X}_n \geq E(X_1) + \varepsilon) \leq e^{-na(\varepsilon)}$$
$$P(\overline{X}_n \leq E(X_1) - \varepsilon) \leq e^{-na(\varepsilon)}.$$

Then extend to the case in which

$$\varphi_{X_n}(u) = \frac{e^{iua}}{1+u^2}.$$

(Try to make no further calculations).

Solution 5. Since $m(t) = m_{X_n}(t) = \varphi_{X_n}(-it) = \frac{1}{1-t^2}$, we have

$$\mu = E(X_n) = m'(0) = 0$$

and

$$g(t) = t\varepsilon - \log\frac{1}{1-t^2},$$

which is maximized at $t^*(\varepsilon) = \frac{\sqrt{1+\varepsilon^2}-1}{\varepsilon}$. So, for the upper tail estimate, it is enough to take $a(\varepsilon) = g(t^*(\varepsilon))$. Since $h(t) = g(t)$ (or equivalently $-X_n$ has the same characteristic function, and so the same distribution, as X_n), the estimate for the lower tail follows.

Finally, if $\varphi_{X_n}(u) = \frac{e^{iua}}{1+u^2}$, then $Y_n := X_n - a$ is such that $\varphi_{Y_n}(u) = \frac{1}{1+u^2}$. Since $E(X_n) = a$,

$$P(\overline{X}_n \geq E(X_1) + \varepsilon) = P(\overline{Y}_n \geq \varepsilon) \leq e^{-na(\varepsilon)}$$

by what seen above. Similarly for the lower tail.

Problem 6. Let $Y_1, Y_2, \ldots, Y_n, Z_1, Z_2, \ldots, Z_n$ be independent random variables such that $Y_i \sim \mathrm{Exp}(1)$ and $Z_i \sim N(0,1)$. Set $X_i := Y_i + Z_i$.

(a) Compute mean, variance and moment generating function of X_i.

(b) For $0 < \varepsilon < 1$, determine $a(\varepsilon) > 0$ such that the following *lower tail Chernoff bound* holds:

$$P(\overline{X}_n \leq 1 - \varepsilon) \leq e^{-na(\varepsilon)}.$$

Hint: use the inequality $\log(1+t) \geq t - \frac{1}{2}t^2$ for $t \geq 0$.

Solution 6. (a) $E(X_i) = E(Y_i) + E(Z_i) = 1$, $Var(X_i) = Var(Y_i) + Var(Z_i) = 2$. Moreover

$$m_{X_i}(t) = m_{Y_i}(t) m_{Z_i}(t) = \frac{1}{1-t} e^{t^2/2}.$$

(b) Set

$$h(t) := -t(E(X_i) - \varepsilon) - \log m_{X_i}(-t) = -t(1-\varepsilon) + \log(1+t) - \frac{t^2}{2} \geq t\varepsilon - t^2.$$

Thus

$$\max(h(t) : t \geq 0) \geq t\varepsilon - t^2|_{t=\varepsilon/2} = \frac{\varepsilon^2}{4}.$$

Thus we can take $a(\varepsilon) = \frac{\varepsilon^2}{4}$.

Problem 7. (a) Let $X_1, X_2, \ldots, X_n$ be i.i.d random variables with distribution $\mathrm{N}(0, \sigma^2)$. Show that for every $\varepsilon > 0$

$$\mathrm{P}(\overline{X}_n > \varepsilon) \leq e^{-n\varepsilon^2/2\sigma^2}.$$

(b) A random variable X with $\mathrm{E}(X) = 0$ is said to be *Subgaussian* for the parameter $\sigma > 0$ if its moment generating function $m_X(t)$ is such that

$$m_X(t) \leq e^{t^2\sigma^2/2}$$

for all $t \in \mathbb{R}$. Show that the inequality in (a) holds if $X_1, X_2, \ldots, X_n$ are i.i.d random variables, and Subgaussian for the parameter $\sigma > 0$.

Solution 7. (a) We use Chernoff bounds. Set $g(t) := \varepsilon t - \log m_{X_1}(t)$. We know that for all t

$$\mathrm{P}(\overline{X}_n > \varepsilon) \leq e^{-ng(t)}.$$

Recalling that

$$m_{X_1}(t) = e^{t^2\sigma^2/2},$$

so

$$g(t) = \varepsilon t - \frac{t^2\sigma^2}{2},$$

whose maximum is attained at $t^* = \frac{\varepsilon}{\sigma^2}$, giving $g(t^*) = \frac{\varepsilon^2}{2\sigma^2}$, from which the desired bound follows.

(b) Just observe that, in this case,

$$g(t) := \varepsilon t - \log m_{X_1}(t) \geq \varepsilon t - \frac{t^2\sigma^2}{2},$$

so $g(t^*) \geq \frac{\varepsilon^2}{2\sigma^2}$ and the bound follows.