Probability Theory Assignment
Due February 1, 2008, at the General Office.

Question 1. (a) Give the definition of the characteristic function $\Psi_X(\theta)$ of a random variable X.

(b) Prove that if X and Y are two independent random variables, then

$$\Psi_{X+Y}(\theta) = \Psi_X(\theta) \cdot \Psi_Y(\theta).$$

It is known that, for any α with $0 \le \alpha \le 2$, for any $c \ge 0$, and $\mu \in \mathbb{R}$ there exist random variables X whose characteristic function is of the form

$$\Psi_X(\theta) = e^{i\mu\theta - c|\theta|^\alpha}.$$

We will say in this case that $X \sim \mathcal{L}(\mu, c; \alpha)$.

(c) Let X_1 and X_2 be independent random variables such that

$$X_1 \sim \mathcal{L}(\mu_1, c_1; \alpha), \quad X_2 \sim \mathcal{L}(\mu_2, c_2; \alpha),$$

that is, they share the same parameter α. Show that

$$X_1 + X_2 \sim \mathcal{L}(\mu_1 + \mu_2, c_1 + c_2; \alpha).$$

(d) Show that if $X_1, \dots, X_n$ are independent with $X_i \sim \mathcal{L}(0, 1; \alpha)$, then

$$\frac{X_1 + \cdots + X_n}{n^{1/\alpha}} \sim \mathcal{L}(0, 1; \alpha).$$

(e) Let $X_1, X_2, \cdots$ be independent and identically distributed, with $X_i \sim \mathcal{L}(\mu, c; \alpha)$. Consider the sequence of averages

$$A_n := \frac{X_1 + \cdots + X_n}{n}, \quad n = 1, 2, \dots \ .$$

Show that, if $\alpha > 1$, then

$$\Psi_{A_n}(\theta) \to e^{i\mu\theta} \quad \text{as } n \to \infty \ .$$

What conclusion can you draw about the sequence $(A_n)_{n \ge 1}$ itself? Discuss what happens when $\alpha = 1$.

Question 2 Let $t \in [0, 1]$ and let $X_1, X_2, \ldots$ be a sequence of independent Bernoulli random variables satisfying

$$\mathbb{P}(X_k = 1) = t \quad \text{and} \quad \mathbb{P}(X_k = 0) = 1 - t,$$

for all $k \geq 1$. Further, let us define the average

$$A_n = \frac{X_1 + X_2 + \cdots + X_n}{n},$$

for positive integer n.

(a) Calculate $\mathbb{E}X_k$ and $\operatorname{Var} X_k$.

(b) Calculate $\mathbb{E}A_n$ and prove that $\operatorname{Var} A_n = t(1-t)/n$.

(c) Hence show that, for any $\delta > 0$,

$$\mathbb{P}\left(|A_n - t| > \delta\right) \leq \frac{t(1-t)}{n\delta^2}.$$

[Hint: Chebyshev's inequality and the Weak Law of Large Numbers.]

Now let $f : [0, 1] \to \mathbb{R}$ be any continuous function. You may assume that f possesses the following properties:

(1) f is *uniformly continuous*: given any $\epsilon > 0$, there exists $\delta > 0$ for which $|f(x) - f(y)| \leq \epsilon$ whenever x and y are points in $[0, 1]$ satisfying $|x - y| \leq \delta$;

(2) f is *bounded*: there exists a non-negative number M for which $|f(x)| \leq M$, for all $x \in [0, 1]$.

(d) The *Bernstein polynomial* $B_n f(t) : [0, 1] \to \mathbb{R}$ is defined by

$$B_n f(t) = \mathbb{E} f(A_n).$$

Prove that

$$B_n f(t) = \sum_{k=0}^{n} f(k/n) \binom{n}{k} t^k (1-t)^{n-k}.$$

(e) Show that

$$|f(t) - B_n f(t)| \leq \mathbb{E}\left|f(A_n) - f(t)\right|.$$

[You may use the fact that $|\mathbb{E}U| \leq \mathbb{E}(|U|)$, for any random variable U.]

(f) Prove that

$$|f(t) - B_n f(t)| \leq \epsilon \mathbb{P}\left(|A_n - t| \leq \delta\right) + 2M\mathbb{P}\left(|A_n - t| > \delta\right).$$

(g) Hence show that

$$|f(t) - B_n f(t)| \leq \epsilon + \frac{2Mt(1-t)}{n\delta^2}.$$

.

(h) Finally, prove that

$$|f(t) - B_n f(t)| \leq 2\epsilon$$

for all $t \in [0, 1]$ if n exceeds $M/(2\delta^2\epsilon)$.

[**Aside:** In other words, the Bernstein polynomials $B_n f(t)$ converge uniformly to $f(t)$. The above proof was found by the great Russian mathematician Bernstein in 1913, at the age of 19, and was the root of an entirely new field of research in which probabilistic methods began to colonize other areas of mathematics. Bernstein polynomials have many important properties and you have almost certainly been using software that makes use of them for many years: they are used to approximate curves in drawing packages and to design the fonts used on this page.]

Question 3 The pupils of two groups of 10000 students, A and B say, compete for a new mathematical scholarship by taking an examination. The marks for pupils in Group A are approximately normally distributed $N(\mu, \sigma^2)$, whilst the marks for pupils in Group B are approximately normally distributed $N(\mu, (1+\delta)^2\sigma^2)$, where δ is positive. The scholarship providers decide that to award scholarships to all students whose marks exceed $\mu + T\sigma$. Further, they provide special "star scholarships" to those whose marks exceed $\mu + (T + 0.5)\sigma$.

(a) Let $A \sim N(\mu, \sigma^2)$ and $B \sim N(\mu, (1+\delta)^2\sigma^2)$. Prove that

$$\mathbb{P}\left(A \geq \mu + T\sigma\right) = 1 - \Phi(T) \quad \text{ and } \quad \mathbb{P}\left(B \geq \mu + T\sigma\right) = 1 - \Phi\left(\frac{T}{1+\delta}\right),$$

where

$$\Phi(x) = \int_{-\infty}^{x} (2\pi)^{-1/2} \exp(-t^2/2)\, dt, \qquad \text{for } x \in \mathbb{R},$$

that is, the $N(0,1)$ cumulative distribution function.

(b) Suppose $\delta = 0.1$ and $T = 2.5$. It transpires that 60 of the 10000 pupils of Group A win scholarships, compared to 110 of the 10000 pupils of Group B. Moreover, 13 Group A pupils win star scholarships, compared to 32 Group B pupils. A truculent local politician is angered by this result, stating "This is clearly discrimination! Since the average marks of the groups are identical, the same proportion should win a scholarship. It's even more egregious discrimination for star scholarships!" Is the politician justified?

(c) The scholarship provider deems any mark lower than $\mu - T\sigma$ to be a failure. The truculent politician makes a new discovery: "I have found that 62 pupils in Group A failed the examination, whilst 120 pupils in Group B failed. This proves that Group B is no better than Group A! How, therefore, can we trust the claim that Group B should be awarded double the proportion of scholarships?" Is the politician justified?

(d) Apply integration by parts to prove that, for $x > 0$,

$$(*) \qquad \int_x^\infty e^{-t^2/2}\, dt = x^{-1} e^{-x^2/2} - R(x).$$

where

$$R(x) = \int_x^\infty t^{-2} e^{-t^2/2}\, dt.$$

Hence prove that

$$\lim_{x \to \infty} \frac{1 - \Phi(x)}{(2\pi)^{-1/2} x^{-1} e^{-x^2/2}} = 1.$$

[Hint: For the last part, divide equation (*) by $x^{-1}\exp(-x^2/2)$ and make the change of variable $t = x + s$ in the integral defining $R(x)$.

(e) Using part (d), prove that

$$\frac{\mathbb{P}\left(\text{Pupil in Group B wins scholarship}\right)}{\mathbb{P}\left(\text{Pupil in Group A wins scholarship}\right)} \approx (1+\delta) e^{\frac{T^2}{2}\left(1-(1+\delta)^{-2}\right)}.$$

What happens as $T \to \infty$?

[**Hint:** The website `www.wessa.net/rwasp_cdfnorm.wasp` provides software to compute Φ. Alternatively, install `R` from `www.r-project.org`.]