PROBABILITY FOR STATISTICS: EXAMINATION SOLUTIONS 2013-14

Q1. (i) *Area of an ellipse.*

The proof given in lectures uses plane polar coordinates to find the area of a circle, and then cartesian polar coordinates to reduce to this case by dilation (or compression) of one axis [see pfsl1, on website]. The following alternative proof uses only one coordinate system.

Let the ellipse E be $x^2/a^2 + y^2/b^2 = 1$. Parametrise the interior E° by $x = ar\cos\theta$, $y = br\sin\theta$ ($\theta \in [0, 2\pi], r \in [0, 1]$). As

$$\partial x/\partial r = a\cos\theta = x/r, \quad \partial y/\partial r = b\sin\theta = y/r,$$

$$\partial x/\partial\theta = -ar\sin\theta = -ay/b, \quad \partial y/\partial\theta = br\cos\theta = bx/a$$

the Jacobian is

$$J = \left| \begin{array}{cc} x/r & -ay/b \\ y/r & bx/a \end{array} \right| = \frac{1}{r}\Big(\frac{b}{a}x^2 + \frac{a}{b}y^2\Big) = \frac{b^2x^2 + a^2y^2}{abr} = \frac{a^2b^2r^2}{abr} = abr.$$

So the area is

$$A = \int\int_{E^\circ} dxdy = \int\int_{E^\circ} Jdrd\theta = \int_0^{2\pi} d\theta . \int_0^1 abrdr = 2\pi . \frac{1}{2}ab = \pi ab.$$

[12]

(ii) *Uniform distribution on subgraphs and densities.*

We are given a density f and so its subgraph $S := \{(x, y) : 0 \le y \le f(x)\}$.
(a) $\Rightarrow$ (b). Uniform distribution over S is w.r.t. the measure $dxdy$ over S. Integrating this over $0 \le y \le f(x)$ to project onto the first coordinate gives the image measure $f(x)dx$, under which the first coordinate, X, has density f: $P(X \in [x, x + dx]) = f(x)dx$. **[6]**
(b) $\Rightarrow$ (a). Conversely, as

$$dxdy = f(x)dx.dy/f(x),$$

if X has density f (as above), and $Y|X = x$ is uniform on $[0, f(x)]$, (X, Y) is uniform on S. **[7]**
(i): seen; (ii): unseen (but should be familiar from simulation – this comes into the rejection method).

Q2 (*Weak Law of Large Numbers, WLLN*).

Random variables X_n *converge in probability* to X if

$$\forall\, \epsilon > 0, \quad P(|X_n - X| > \epsilon) \to 0 \qquad (n \to \infty). \qquad [\mathbf{2}]$$

They converge to X *in distribution* if

$$F_n(x) := P(X_n \le x) \to F(x) := P(X \le x) \qquad (n \to \infty),$$

at all continuity points x of F. **[2]**

(a) Convergence in probability implies convergence in distribution, but not conversely in general. **[1]**

(b) The converse holds (so the two are equivalent) if the limit X is constant. **[1]**

We need the following properties of the characteristic function (CF):

(i) (Lévy's convergence theorem). Convergence in distribution of random variables is equivalent to convergence of CFs (uniformly on compacta). **[2]**

(ii). The CF of an independent sum is the product of the CFs. **[2]**

(iii). If the random variable has k moments (finite), the CF can be expanded as far as the t^k term with negligible error term $o(t^k)$ for small t. **[2]**

Recall that (for x real and $z_n \to z$ complex)

$$(1 + \frac{x}{n})^n \to e^x, \qquad (1 + \frac{z_n}{n})^n \to e^z \qquad (n \to \infty). \qquad (*)\ [\mathbf{2}]$$

Theorem (Weak Law of Large Numbers, WLLN). **[3]**

If X_i are iid with mean μ,

$$\frac{1}{n}\sum_1^n X_k \to \mu \qquad (n \to \infty) \qquad \text{in probability.}$$

Proof. If the X_k have CF $\phi(t)$, then as the mean μ exists $\phi(t) = 1 + i\mu t + o(t)$ as $t \to 0$, by (iii). So by (ii) $(X_1 + \ldots + X_n)/n$ has CF

$$E \exp\{it(X_1 + \ldots + X_n)/n\} = [\phi(t/n)]^n = [1 + \frac{i\mu t}{n} + o(1/n)]^n,$$

for fixed t and $n \to \infty$. By $(*)$, the RHS has limit $e^{i\mu t}$ as $n \to \infty$. But $e^{i\mu t}$ is the CF of the constant μ. So by Lévy's continuity theorem (i),

$$(X_1 + \ldots + X_n)/n \to \mu \qquad (n \to \infty) \qquad \text{in distribution.}$$

Since the limit μ is constant, this and (b) give

$$(X_1 + \ldots + X_n)/n \to \mu \qquad (n \to \infty) \qquad \text{in probability. } // \qquad [\mathbf{8}]$$

Seen – lectures.

Q3 (*Compound Poisson processes*).
(i) Let jumps $\{X_1, \cdots, X_n, \cdots\}$ arrive at the epochs of a Poisson process with rate λ [claims, in the insurance context used in lectures – call them claims below, for definiteness]. Then the number $N(t)$ of claims in the time-interval $[0,t]$ is Poisson $P(\lambda t)$. If the claims are iid with mean μ and CF $\phi(u)$, then the claim total at time t is $S(t) := X_1 + \cdots X_{N(t)}$, with CF

$$\psi(u) := E[\exp\{iuS(t)\}] = E[\exp\{iu(X_1 + \cdots + X_{N(t)})]$$

$$= \sum_{n=0}^{\infty} E[\exp\{iu(X_1 + \cdots + X_{N(t)})\}|N(t) = n]P(N(t) = n)$$

$$= \sum E[\exp\{iu(X_1 + \cdots + X_n)\}].e^{-\lambda t}(\lambda t)^n/n! = e^{-\lambda t}\exp\{\lambda t\phi(u)\}. \qquad \textbf{[7]}$$

From $\phi(u) := E[e^{iuX}]$, $\phi'(u) = iE[Xe^{iuX}]$, $\phi''(u) = -E[X^2 e^{iuX}]$, so $\phi'(0) = iE[X] = i\mu$, say, $\phi''(0) = -E[X^2]$, and similarly
$\psi'(0) = iE[S(t)]$, $\psi''(0) = -E[S(t)^2]$. **[2]**
So differentiating,

$$\psi'(u) = \phi'(u).\lambda t.\psi(u); \qquad \psi'(0) = \lambda t.\phi'(0);$$

$$\psi''(u) = \lambda t.\phi''(u).\psi(u) + \lambda t.\phi'(u), \psi'(u) = \lambda t\phi''(u)\psi(u) + (\lambda t)^2[\phi'(u)]^2\psi(u) :$$

$$\psi''(0) = \lambda t\phi''(0) + (\lambda t)^2[\phi'(0)]^2. \qquad \textbf{[2]}$$

$$E[S(t)] = \lambda tE[X] = \lambda t\mu; \qquad \textbf{[3]}$$

$$var[S(t)] = -\psi''(0) = [\psi'(0)]^2 = -\lambda t\phi''(0) + (\lambda t)^2\mu^2 - (\lambda t)^2\mu^2 = \lambda tE[X^2]. \qquad \textbf{[4]}$$

(ii) If the claim times are T_n, the number $N(t)$ of claims to date is n on $[T_n, T_{n+1})$, jumping to $n+1$ at T_{n+1}. So $T_{N(t)} \leq t < T_{N(t)+1}$. So

$$\frac{T_{N(t)}}{N(t)} \leq \frac{t}{N(t)} < \frac{T_{N(t)+1}}{N(t)+1}.\frac{N(t)+1}{N(t)}.$$

Let $t \to \infty$: $N(t) \to \infty$ also, through *integer* values, n say. So the left (a sample mean – an average of n iid rvs, each Poisson $P(\lambda)$ with mean λ) tends to λ a.s., by the Strong Law of Large Numbers. Similarly, so does the first factor on the right, while the second factor on the right tends to 1 a.s. Combining, the inequality above gives

$$\frac{t}{N(t)} \to \lambda \qquad a.s. : \qquad \frac{N(t)}{t} \to 1/\lambda \qquad a.s. \qquad \textbf{[7]}$$

(i): Method, results, and most proofs, seen; (ii): unseen.

Q4 (*Finite Markov chains*).

A state i in a Markov chain is *transient* if the chain spends only finitely long in i (a.s.), *recurrent* (or *persistent*) otherwise. **[1,1]**

The *mean recurrence time* of a state i is the expectation of the time T_i of first return to i, starting at i. **[1]**

A recurrent state (one to which the chain returns infinitely often (i.o.), a.s.) is *positive* if the mean recurrence time is finite, *null* otherwise. **[1,1]**

Theorem. For a finite Markov chain, it is impossible for all states to be transient: a finite chain must contain at least one persistent state.

Proof. If the state-space is $\{1, \cdots, N\}$, for each i and each n

$$1 = \sum_{j=1}^{N} p_{ij}(n). \qquad (a)$$

Let $n \to \infty$: if j is transient, the total expected time in it is finite: $\sum_n p_{ij}(n) < \infty$. So

$$p_{ij}(n) \to 0 \qquad (n \to \infty). \qquad (b)$$

Were *all* states transient, letting $n \to \infty$ in (a) and using (b) would give the contradiction $1 = 0$. So not all states in a finite chain can be transient. // **[7]**

Theorem. A recurrent state j in a finite chain is positive (= non-null).

Proof. If the finite chain has state-space $\{1, \cdots, N\}$, assume there is a null state. Let C be the equivalence class containing it. Since C is closed, we can consider the subchain induced on C. Then

$$1 = \sum_{k \in C} p_{ik}(n) \qquad \text{(finite sum)}.$$

Let $n \to \infty$: each $p_{ik}(n) \to 0$, so RHS $\to 0$, giving $1 = 0$. This contradiction gives the non-existence of null states in a finite chain. // **[7]**

All states may be :

(i) transient. Trivial example: $\mathbb{Z}$, moving to the right at each step. **[2]**
(ii) positive recurrent. Trivial example: $\mathbb{Z}$, with each state a trap. **[2]**
(iii) null recurrent. Example: simple random walk on $\mathbb{Z}$. **[2]**
Seen – lectures.

N. H. Bingham