pfsmockexamsoln.tex 13.8.2012

# PROBABILITY FOR STATISTICS: MOCK EXAM SOLUTIONS, 2012

Q1. The *chi-square distribution* with $n$ *degrees of freedom*, $\chi^2(n)$, is defined as the distribution of $X_1^2 + \ldots + X_n^2$, where $X_i$ are iid $N(0,1)$.
(i) For $n = 1$, the mean is 1, because a $\chi^2(1)$ is the square of a standard normal, and a standard normal has mean 0 and variance 1. The variance is 2, because the fourth moment of a standard normal $X$ is 3, and

$$var(X^2) = E[(X^2)^2] - [E(X^2)]^2 = 3 - 1 = 2.$$

For general $n$, the mean is $n$ because means add, and the variance is $2n$ because variances add over independent summands.
(ii) For $X$ standard normal, the MGF of its square $X^2$ is

$$M(t) := \int e^{tx^2}\phi(x)dx = \frac{1}{\sqrt{2\pi}}\int_0^\infty e^{tx^2}.e^{-\frac{1}{2}x^2}dx.$$

We see that the integral converges only for $t < \frac{1}{2}$, when (normal integral) it is $1/\sqrt{(1-2t)}$:

$$M(t) = 1/\sqrt{1-2t} \qquad (t < \frac{1}{2}) \qquad \text{for } X\ N(0,1).$$

So by definition of $\chi^2(n)$, the MGF of a $\chi^2(n)$ is

$$M(t) = 1/(1-2t)^{\frac{1}{2}n} \qquad (t < \frac{1}{2}) \qquad \text{for } X\ \chi^2(n).$$

(iii) First, the required density $f(.)$ is a density, as $f \geq 0$ and $\int f = 1$:

$$\int f(x)dx = \frac{1}{2^{\frac{1}{2}n}\Gamma(\frac{1}{2}n)}.\int_0^\infty x^{\frac{1}{2}n-1}\exp(-\frac{1}{2}x)dx = \frac{1}{\Gamma(\frac{1}{2}n)}.\int_0^\infty u^{\frac{1}{2}n-1}\exp(-u)du = 1$$

($u := \frac{1}{2}x$), by definition of the Gamma function. Its MGF is

$$M(t) = \frac{1}{2^{\frac{1}{2}n}\Gamma(\frac{1}{2}n)}.\int_0^\infty e^{tx}.x^{\frac{1}{2}n-1}\exp(-\frac{1}{2}x)dx = \frac{1}{2^{\frac{1}{2}n}\Gamma(\frac{1}{2}n)}.\int_0^\infty x^{\frac{1}{2}n-1}\exp(-\frac{1}{2}x(1-2t))dx.$$

Substitute $u := \frac{1}{2}x(1-2t)$ in the integral. One obtains

$$M(t) = (1-2t)^{-\frac{1}{2}n}.\frac{1}{\Gamma(\frac{1}{2}n)}.\int_0^\infty u^{\frac{1}{2}n-1}e^{-u}du = (1-2t)^{-\frac{1}{2}n}.$$

So it has the required MGF, so is the required density. //

Q2. An $n$-vector $X$ has an *$n$-variate normal* (or *Gaussian*) distribution iff $a^T X$ is univariate normal for all constant $n$-vectors $a$.

**Theorem**. If $X$ is $n$-variate normal with mean $\mu$ and covariance matrix $\Sigma$, its CF is

$$\phi(t) := E[\exp\{it^T X\} = \exp\{it^T \mu - \frac{1}{2} t^T \Sigma t\}.$$

*Proof*. By the Proposition, $Y := t^T X$ has mean $t^T \mu$ and variance $t^T \Sigma t$. By definition of multinormality, $Y = t^T X$ is univariate normal. So $Y$ is $N(t^T \mu, t^T \Sigma t)$. So $Y$ has CF

$$\phi_Y(s) := E[\exp\{isY\}] = \exp\{ist^T \mu - \frac{1}{2} t^T \Sigma t\}.$$

But $E[(e^{isY})] = E[\exp\{ist^T X\}]$, so taking $s = 1$ (as in the proof of the Cramér-Wold device),

$$E[\exp\{it^T X\} = \exp\{it^T \mu - \frac{1}{2} t^T \Sigma t\},$$

giving the CF of $X$ as required. //

**Theorem**. Linear forms $Ax$ and $Bx$ with $x \sim N(\mu, \Sigma)$ are independent iff

$$A \Sigma B^T = 0.$$

In particular, if $A$, $B$ are symmetric and $\Sigma = \sigma^2 I$, they are independent iff

$$AB = 0.$$

*Proof*. The joint CF is

$$\phi(u, v) := E[\exp\{iu^T A + iv^T Bx\}] = E[\exp\{i(A^T u + B^T v)^T x\}].$$

This is the CF of $x$ at argument $t = A^T u + B^T v$, so

$$\phi(u, v) = \exp\{i(u^T A + v^T B)\mu - \frac{1}{2}[u^T A \Sigma A^T u + u^T A \Sigma B^T v + v^T B \Sigma A^T u + v^T B \Sigma B^T v]\}.$$

This factorises into a product of a function of $u$ and a function of $v$ iff the two cross-terms in $u$ and $v$ vanish, that is, iff $A \Sigma B^T = 0$ and $B \Sigma A^T = 0$; by symmetry of $\Sigma$, the two are equivalent. //

Q3. As $f \geq 0$ and

$$\begin{aligned}
\int f &= \frac{1}{\sqrt{2\pi}} \int_0^\infty x^{-3/2} \exp\Big(-\frac{1}{2x}\Big) dx \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty u^{\frac{1}{2}} \exp(-\frac{1}{2}u) du \qquad (u := 1/x) \\
&= \frac{1}{\sqrt{2\pi}}.\sqrt{2} \int_0^\infty v^{\frac{1}{2}} \exp(-v) dv \qquad (v := \frac{1}{2}u) \\
&= \frac{1}{\sqrt{\pi}} \Gamma(\frac{1}{2}) = 1 \qquad (\Gamma(\frac{1}{2}) = \sqrt{\pi}),
\end{aligned}$$

$f$ is a density. Its Laplace-Stieltjes transform is

$$\phi(s) = \frac{1}{\sqrt{2\pi}}. \int_0^\infty \exp(-sx - \frac{1}{2x}). \frac{dx}{x^{3/2}}.$$

Differentiate under the integral sign (as we may, the integrand being monotone in $s$ – we quote this):

$$\phi'(s) = -\frac{1}{\sqrt{2\pi}}. \int_0^\infty \exp(-sx - \frac{1}{2x}). \frac{dx}{\sqrt{x}}.$$

The change of variable suggested interchanges the two terms in the exponential. It reverses the limits, and

$$\frac{dx}{\sqrt{x}} = -\frac{1}{\sqrt{2s}}. \frac{du}{u^{3/2}}.$$

This gives

$$\phi'(s) = -\frac{1}{\sqrt{2s}}.\phi(s): \qquad \frac{\phi'(s)}{\phi(s)} = -\frac{1}{\sqrt{2s}}.$$

Integrate: $\log \phi(s) = -\sqrt{2s} + c$, $\phi(s) = ce^{-\sqrt{2s}}$. But $\phi(0) = \int f = 1$, so $c = 1$ and $\phi(s) = e^{-\sqrt{2s}}$. //

Q4. The number of subsets of size $d$ of a set of size $2d$ is $\binom{2d}{d}$. With $d$ white balls and $d$ black balls; the $i$ white balls can be chosen in $\binom{d}{i}$, the remaining $d-i$ black balls in $\binom{d}{d-1} = \binom{d}{i}$ ways. So $\binom{2d}{d} = \sum_0^d \binom{d}{i}^2$ counts the total number of ways by how many white balls are chosen. So $\pi_i$ defines a pr. distribution (the *hypergeometric distribution* $HG(d)$).

$$\binom{d}{j-1}(d-j+1) = \frac{d!}{(j-1)!(d-j+1)!}.(d-j+1) = \frac{d!}{(j-1)!(d-j)!} = \binom{d}{j}.j, \qquad (i)$$

$$\binom{d}{j+1}(j+1) = \frac{d!}{(j+1)!(d-j-1)!}(j+1) = \frac{d!}{j!(d-j-1)!} = \binom{d}{j}.(d-j). \qquad (ii)$$

So by (i) and (ii),

$$\begin{aligned}(\pi P)_j &= \sum_i \pi_i p_{ij} = \pi_{j-1}p_{j-1,j} + \pi_{j+1}p_{j+1,j} \\ &= \frac{1}{d^2\binom{2d}{d}}\Big[\binom{d}{j-1}^2(d-j+1)^2 + \binom{d}{j}^2.2j(d-j) + \binom{d}{j+1}^2(j+1)^2\Big] \\ &= \frac{\binom{d}{j}^2}{d^2\binom{2d}{d}}\{j^2 + 2d(d-j) + (d-j)^2\} = \frac{\binom{d}{j}^2}{\binom{2d}{d}} = \pi_j.\end{aligned}$$

So $\pi P = \pi$, and $\pi$ is invariant, as required. //

$$\pi_i p_{i,i+1} = \binom{2d}{d}^{-1}\Big(\binom{d}{i}.\frac{d-i}{d}\Big)^2 = \binom{2d}{d}^{-1}\Big(\frac{d!}{(d-i)!i!}.\frac{d-i}{d}\Big)^2 = \binom{2d}{d}^{-1}\binom{d-1}{i}^2,$$

and similarly

$$\pi_{i+1}p_{i+1,i} = \binom{2d}{d}^{-1}\binom{d-1}{i}^2,$$

proving detailed balance, and so reversibility. Then

$$\pi_i = \pi_0\Big(\frac{1}{\frac{1}{d}}.\frac{1-\frac{1}{d}}{\frac{2}{d}}.\dots.\frac{1-\frac{i-1}{d}}{\frac{i}{d}}\Big)^2 = \Big(\pi_0.\frac{d(d-1)\dots(d-i+1)}{1.2\dots.i}\Big)^2 = \pi_0\binom{d}{i}^2.$$

Then $\sum_i \pi_i = 1$ gives

$$\pi_0\sum_i\binom{d}{i}^2 = \pi_0.\binom{2d}{d} = 1, \quad \pi_0 = \binom{2d}{d}^{-1}, \quad \pi_i = \binom{d}{i}^2\Big/\binom{2d}{d}, \quad \pi = HG(d).$$