ENGINEERING TRIPOS PART IB

Friday 9 June 2006 9 to 11

Paper 7

MATHEMATICAL METHODS

Solutions

SECTION A

1 (a) Imagine the volume V to be composed of a large number of elemental volumes δV of any shape. The surfaces of these elemental volumes are mostly composed of common interfaces between adjoining elements, but each element that is at the edge of the volume will include a small part of the external surface S as part of its own surface. For each element δV, the net flux of $\mathbf{F}$ is given by $(\nabla \cdot \mathbf{F})\delta V$. In the interior of the volume, the flux out of one element is equal to the flux into an adjacent element across the common interface. All these contributions to the volume integral cancel. It is only at the surface S that the contributions survive and sum to give the net flux from the whole volume:

$$\int_V (\nabla \cdot \mathbf{F}) dV = \oint_S \mathbf{F} \cdot dA$$

This argument is required in full. It is not sufficient to sum the definition of divergence over a volume V without noting that contributions from adjacent surfaces of interior cells cancel.

(b) The volume of air within the device is:

$$\begin{aligned} V = \int_0^{H_2} \pi r^2 \ dz &= \int_0^{H_1} \pi R^2 \ dz + \int_{H_1}^{H_2} \pi \left(\frac{RH_1}{z} \right)^2 \ dz \\ &= \pi R^2 H_1 + \pi R^2 H_1^2 \left[-\frac{1}{z} \right]_{H_1}^{H_2} \\ &= \pi R^2 H_1 + \pi R^2 H_1^2 \left(\frac{1}{H_1} - \frac{1}{H_2} \right) \\ &= \pi R^2 H_1 \left(2 - \frac{H_1}{H_2} \right) \end{aligned}$$

(c) Gauss' theorem relates the divergence within the device to the flux across its surface:

$$\begin{aligned} \iiint_V \nabla \cdot \mathbf{u} \ dV = \iint_S \mathbf{u} \cdot \mathbf{n} \ \ dS &= \iint_{turbine} \mathbf{u} \cdot \mathbf{n} \ dS + \iint_{net} \mathbf{u} \cdot \mathbf{n} \ dS \\ &= \iint_{turbine} U_2 \hat{\mathbf{e}}_\mathbf{z} \cdot \hat{\mathbf{e}}_\mathbf{z} \ dS + \iint_{net} -U_1 \hat{\mathbf{e}}_\mathbf{r} \cdot \hat{\mathbf{e}}_\mathbf{r} \ dS \\ &= U_2 \pi \left(\frac{RH_1}{H_2} \right)^2 - 2U_1 \pi R H_1 \\ &= \pi R \frac{H_1}{H_2} \left(\frac{RH_1}{H_2} U_2 - 2H_2 U_1 \right) \end{aligned}$$

It is important to state explicitly that $\nabla \cdot \mathbf{u}$ can be taken out of the integral on the LHS because it is uniform within the device:

$$
\begin{aligned}
\nabla \cdot \mathbf{u} \int\int\int_V dV &= \pi R \frac{H_1}{H_2}\left(\frac{RH_1}{H_2}U_2 - 2H_2U_1\right) \\
\Rightarrow \quad \nabla \cdot \mathbf{u} &= \frac{1}{V}\pi R \frac{H_1}{H_2}\left(\frac{RH_1}{H_2}U_2 - 2H_2U_1\right) \\
&= \frac{\pi R \frac{H_1}{H_2}\left(\frac{RH_1}{H_2}U_2 - 2H_2U_1\right)}{\pi R^2 H_1\left(2 - \frac{H_1}{H_2}\right)} \\
&= \frac{RH_1U_2 - 2H_2^2U_1}{R(2H_2^2 - H_1H_2)}
\end{aligned}
$$

(d) By Stokes' theorem, the flux of $\nabla \cdot \mathbf{u}$ through a cross-section at $z = H_1/2$ is equal to the line integral of $\mathbf{u}$ around the cross-section's perimeter, at R:

$$
\begin{aligned}
\int\int_S (\nabla \times \mathbf{u}) \cdot \mathbf{n} \; dS &= \oint \mathbf{u} \cdot d\mathbf{l} \\
&= \int_0^{2\pi} -U_1 \hat{\mathbf{e}}_\mathbf{r} \cdot 2\pi R \hat{\mathbf{e}}_\theta \; d\theta \\
&= 0
\end{aligned}
$$

because $\hat{\mathbf{e}}_\mathbf{r}$ and $\hat{\mathbf{e}}_\theta$ are perpendicular.

2 (a) By inspection, the scalar potential ϕ is equal to $fyz + C$, where C is a constant. If this is found first, there is no need to show explicitly that $\nabla \times \mathbf{B} = 0$. An explicit calculation would result in:

$$\nabla \times \mathbf{B} = \begin{bmatrix} f - f \\ y\frac{df}{dx} - y\frac{df}{dx} \\ z\frac{df}{dx} - z\frac{df}{dx} \end{bmatrix} = 0$$

(b) Because $\mathbf{B}$ is a conservative vector field,

$$\begin{aligned}
\int_{(-1,1,1)}^{(1,1,1)} \mathbf{B} \cdot d\mathbf{l} &= \phi_{(1,1,1)} - \phi_{(-1,1,1)} \\
&= [fyz]_{(1,1,1)} - [fyz]_{(-1,1,1)} \\
&= \sin(\pi/2) - \sin(-\pi/2) = 2 \qquad \text{when} \qquad f(x) = \sin(\pi x/2) \\
&= \cos(\pi/2) - \cos(-\pi/2) = 0 \qquad \text{when} \qquad f(x) = \cos(\pi x/2)
\end{aligned}$$

(c) Most candidates simply wrote down $\nabla \cdot \mathbf{B} = 0$ but this only leads to one of the possible conditions. By Gauss' theorem:

$$\begin{aligned}
0 = \int\int_S \mathbf{B} \cdot \mathbf{n}\ dS &= \int\int\int_V (\nabla \cdot \mathbf{B})\ dV \\
&= \int\int\int_V \left(yz\frac{d^2 f}{dx^2} \right)\ dV
\end{aligned}$$

The most obvious condition is that $d^2 f/dx^2 = 0$ throughout the sphere. Another is that $\int (d^2 f/dx^2)\, dx = 0$, which leads to the condition that df/dx is symmetric about the plane $x = 0$.

(d)

$$\begin{aligned}
\mathbf{C} &= \psi \mathbf{B} \\
0 = \nabla \times \mathbf{C} &= \nabla \times (\psi \mathbf{B}) \\
&= \psi(\nabla \times \mathbf{B}) + \nabla \psi \times \mathbf{B} \\
&= \nabla \psi \times \mathbf{B} \\
\Rightarrow \quad 0 &= \nabla \psi \times \nabla \phi
\end{aligned}$$

If $\psi = g(\phi)$, where g is any differentiable function of ϕ then:

$$\begin{aligned}
\nabla \psi &= \frac{dg}{d\phi} \nabla \phi \\
\text{hence} \qquad \nabla \psi \times \nabla \phi &= \frac{dg}{d\phi} \nabla \phi \times \nabla \phi = 0
\end{aligned}$$

The chain rule must used in the final part. It is not sufficient to state, without proof, that $\nabla\phi$ and ∇g are parallel because, although true, this can be seen by reverse engineering the solution.

3 (a) If $y(x,t) = X(x)T(t)$ then $c^2 X'''' T + X\ddot{T} = 0$ and therefore:

$$\frac{X''''}{X} = -\frac{1}{c^2}\frac{\ddot{T}}{T} = \pm k^4$$

There are three possible sets of governing ordinary differential equations for real k, although most candidates only wrote down the first set:

(i) $X'''' = k^4 X$ and $\ddot{T} = -c^2 k^4 T$

(ii) $X'''' = -k^4 X$ and $\ddot{T} = c^2 k^4 T$

(iii) $X'''' = 0$ and $\ddot{T} = 0$

By inspection, (i) must be the correct set of ODEs because it has a harmonic solution in time. By contrast, (ii) has exponentially decaying (or growing) solutions in time and (iii) has no oscillating term in time.

The relations for X and T can be substituted into the ODEs (i) above to confirm that they satisfy the equations. The relation between ω, k and c is $k = \sqrt{\omega/c}$.

(b)

$$\frac{\partial^2 y}{\partial x^2} = Tk^2\left[-A\cos kx - B\sin kx + C\cosh kx + D\sinh kx\right]$$

At $x = 0$, the constraint that $y = 0$ implies that $A + C = 0$.
At $x = 0$, the constraint that $\partial^2 y/\partial x^2 = 0$ implies that $-A + C = 0$.
These can only be satisfied when $A = C = 0$.
At $x = L$, the constraint that $y = 0$ implies that $B\sin kL + D\sinh kL = 0$.
At $x = L$, the constraint that $\partial^2 y/\partial x^2 = 0$ implies that $-B\sin kL + D\sinh kL = 0$.
This seemed to be the most challenging part of the question. Adding these two equations together gives $D = 0$. Then either equation gives $k = n\pi/L$, where n is an integer (unless $B = 0$, which is a trivial solution).
The constraint of zero initial velocity gives:

$$\frac{\partial y}{\partial t} = X\omega\left[-P\sin\omega t + Q\cos\omega t\right] = 0$$

This implies that $Q = 0$.

The solution is:

$$y(x,t) = Y_n \sin\left(\frac{n\pi x}{L}\right)\cos(\omega t)$$

where Y_n is some constant for that value of n. Around two thirds of the candidates forgot that this needs to be summed over all n (i.e. over all mode shapes), noting that ω is also a function of n:

$$y(x,t) = \sum_n Y_n \sin\left(\frac{n\pi x}{L}\right) \cos\left(\left[\frac{n\pi}{L}\right]^2 ct\right)$$

4 (a) A real symmetric matrix A will have real eigenvalues and real, orthogonal eigenvectors. Therefore, a valid value for a is 2.

(b) $\det A = 1(3-1) - a(2-0) = 0$ for A singular. $\implies 2 - 2a = 0$. Which tells us that $a = 1$.

For $a = 1$ our matrix becomes

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

and eigenvalues are found by solving the equation $\det(A - \lambda I) = 0$.

$$\implies \begin{vmatrix} 1-\lambda & a & 0 \\ 2 & 3-\lambda & 1 \\ 0 & 1 & 1-\lambda \end{vmatrix} = (1-\lambda)[(3-\lambda)(1-\lambda)-1] - 2(1-\lambda) = 0$$

which gives

$$(1-\lambda)[\lambda^2 - 4\lambda + 3 - 1 - 2] = (1-\lambda)(\lambda^2 - 4\lambda) = \lambda(1-\lambda)(\lambda-4) = 0$$

Giving $\lambda = 0, 1, 4$.

The largest eigenvalue is therefore 4 and we find the eigenvector corresponding to this as follows;

For $\lambda = 4$

$$\begin{bmatrix} -3 & 1 & 0 \\ 2 & -1 & 1 \\ 0 & 1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \implies 3x_1 = x_2, \;\; 2x_1 - x_2 + x_3 = 0, \;\; x_2 = 3x_3$$

Therefore the normalised eigenvector corresponding to $\lambda = 4$ is

$$\pm \frac{1}{\sqrt{11}} [1, 3, 1]^T$$

(c) For the matrix A we find the eigenvalues and vectors in the usual way:

$$\det(A - \lambda I) = 0$$

$$\Longrightarrow \begin{vmatrix} 1-\lambda & a & 0 \\ 2 & 3-\lambda & 1 \\ 0 & 1 & 1-\lambda \end{vmatrix} = 0$$

$$\Longrightarrow (1-\lambda)[(3-\lambda)(1-\lambda)-1]-2a(1-\lambda)=0$$

$$\Longrightarrow (1-\lambda)[\lambda^2-4\lambda+2(1-a)]=0$$

$$\Longrightarrow \lambda=1 \quad \text{and} \quad \lambda=\frac{4\pm\sqrt{(16-8(1-a))}}{2}=2\pm\sqrt{2(1+a)}$$

Thus for there to be only 2 distinct eigenvalues, we must have 2 repeated roots.

This can occur if $16-8(1-a)=0$ ie $a=-1$. For $a=-1$, eigenvalues are therefore 1,2,2;

OR

if $2\pm\sqrt{2(1+a)}=1$. Clearly we need the -ve value, so that $2-\sqrt{2(1+a)}=1$ or $(1+a)=1/2$. Therefore $a=-1/2$ and in this case eigenvalues are 1,1,3.

For $a=-1$, we find the eigenvectors corresponding to these eigenvalues as follows:

$$(A-I)\mathbf{x}=0 \quad \text{and} \quad (A-2I)\mathbf{x}=0$$

For $\lambda=1$

$$\begin{bmatrix} 0 & -1 & 0 \\ 2 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Longrightarrow \quad -x_2=0, \quad 2x_1+2x_2+x_3=0, \quad x_3=-2x_1$$

Therefore the normalised eigenvector corresponding to $\lambda=1$ is ,

$$\pm\frac{1}{\sqrt{5}}[1,0-2]^T$$

For $\lambda=2$

$$\begin{bmatrix} -1 & -1 & 0 \\ 2 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Longrightarrow \quad x_1=-x_2, \quad 2x_1+x_2+x_3=0, \quad x_2=x_3$$

Therefore the normalised eigenvector corresponding to $\lambda=2$ is

$$\pm\frac{1}{\sqrt{3}}[1,-1,-1]^T$$

(TURN OVER for continuation of Question 4

For $a = -1/2$ we similarly obtain another set of eigenvectors:

the normalised eigenvector corresponding to $\lambda = 1$ is

$$\pm\frac{1}{\sqrt{5}}[1,0,-2]^T$$

the normalised eigenvector corresponding to $\lambda = 3$ is

$$\pm\frac{1}{\sqrt{21}}[-1,4,2]^T$$

Note here that although we have a repeated eigenvalue (in both cases), we only have one valid eigenvector corresponding to that eigenvalue. A is therefore known as a *defective* matrix and is therefore not diagonalizable (decomposition process to perform diagonalisation would fail).

(d) Suppose we take some initial vector, $\mathbf{u}_0$. If the matrix is not defective, the distinct eigenvectors will span the image space, so our $\mathbf{u}_0$ can be expanded in terms of the eigenvectors

$$\mathbf{u}_0 = \alpha_0\mathbf{e}_0 + \alpha_1\mathbf{e}_1 + \alpha_4\mathbf{e}_4$$

where $\mathbf{e}_i$ corresponds to the eigenvalue $\lambda = i$. Thus, when we apply A to this equation multiple times we get

$$\begin{aligned} A\mathbf{u}_0 &= 0\alpha_0\mathbf{e}_0 + 1\alpha_1\mathbf{e}_1 + 4\alpha_4\mathbf{e}_4 \\ A^2\mathbf{u}_0 &= 1^2\alpha_1\mathbf{e}_1 + 4^2\alpha_4\mathbf{e}_4 \\ \\ A^n\mathbf{u}_0 &= 1^n\alpha_1\mathbf{e}_1 + 4^n\alpha_4\mathbf{e}_4 \end{aligned}$$

For large n the 4^n clearly dominates and so we have that

$$A^n\mathbf{u}_0 \approx 4^n\alpha_4\mathbf{e}_4$$

$\implies \mathbf{e}_4 \propto A^n\mathbf{u}_0$.

Thus, we are able to obtain an estimate for the eigenvector corresponding to the largest eigenvalue provided our initial guess does have a component in the direction of this eigenvector.

For A as given in part (b), we can apply this method using the initial value given:

$$\mathbf{u}_1 = A\mathbf{u}_0 = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbf{u}_2 = A\mathbf{u}_1 = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}$$

$$\mathbf{u}_3 = A\mathbf{u}_2 = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 16 \\ 6 \end{bmatrix}$$

$$\mathbf{u}_4 = A\mathbf{u}_3 = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 16 \\ 6 \end{bmatrix} = \begin{bmatrix} 21 \\ 64 \\ 22 \end{bmatrix}$$

If we normalise this 4th iteration we obtain an estimate of the eigenvector as

$$\mathbf{e}_4 \approx [0.296,\ 0.903,\ 0.311\,]$$

this is a reasonable approximation when we compare it to the true eigenvector given in (b)

$$\mathbf{e}_4 = [0.302,\ 0.905,\ 0.302\,]$$

with each element of the unit eigenvector being accurate to 1 d.p.

The convergence factor is the ratio $|\lambda_1|/|\lambda_4|$, ie the magnitude of the ratio of the 2nd highest eigenvalue to the highest eigenvalue – effectively the error is reduced by this factor at each iteration (leading term error). In this case the convergence factor is $1/4$, so we expect fairly rapid convergence.

5 (a) To perform LU decomposition do Gaussian elimination on rows of A: first add row 1 to row 2, subtract row 1 from row 3:

$$A = \begin{bmatrix} 2 & -1 & 0 & 3 \\ -2 & 2 & 1 & -4 \\ 2 & -2 & -1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ & 1 & 0 \\ & & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 & 3 \\ 0 & 1 & 1 & -1 \\ 0 & -1 & -1 & 1 \end{bmatrix}$$

We can then see that $l_{21} = -1$ and $l_{31} = 1$. Next we add the new row 2 to row 3.

$$= \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 & 3 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

So that $l_{32} = -1$. The LU decomposition is therefore:

$$LU = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 & 3 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

For the 4 fundamental subspaces we have:

(i) Column Space: given by columns of independent columns of L (which are first 2 as 3rd row of U is zero):

$$= \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

or, we can take the columns of A corresponding to the columns of U with non-zero pivots:

$$= \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ -2 \end{bmatrix}$$

(ii) Row Space: given by the row space of U which is clearly the first 2 rows:

$$[2, -1, 0, 3]^T, \quad [0, 1, 1, -1]^T$$

(iii) Null space of A is the null space of U, therefore

$$\begin{bmatrix} 2 & -1 & 0 & 3 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Have 2 equations in 4 unknowns, therefore need 2 free variables, choose x_3,x_4.

$2x_1 - x_2 + 3x_4 = 0, x_2 + x_3 - x_4 = 0.$

$\implies \; x_1 = -\frac{1}{2}x_3 - x_4, \; x_2 = -x_3 + x_4$ so that

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -1/2 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Which give us the above 2 vectors as a basis for the null space. If we choose different free variables we get different vectors, so anything which is a linear combination of the above is allowed.

(iv) The left null space can be found by either finding the null space of A^T or by taking the last row of L^{-1}, or by taking the cross product of the column space vectors (this is the easiest method). By inspection it is easy to see that L^{-1} is given by

$$L^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

So that the basis for the LNS is $[0,1,1]^T$.

If we do it by finding the null space of A^T we can either straightforwardly gaussian eliminate or do another LU decomposition on A^T.

(b) (i) To find the value of a we integrate the pdf

$$\int_0^2 \int_0^2 p(x,y)\,dx\,dy = \alpha \int_{y=0}^{y=2} \left[\frac{1}{2}x^2y + 2x\right]_0^2 dy \quad (1)$$

$$= \alpha \int_0^2 (2y+4)dy \quad (2)$$

$$= \alpha \left[y^2 + 4y\right]_0^2 = 12\alpha \quad (3)$$

Therefore $\alpha = 1/12$.

(ii)

$$E[X] = \int_0^2 \int_0^2 xp(x,y)\,dx\,dy \quad (4)$$

$$= \alpha \int_0^2 \int_0^2 (x^2y + 2x)\,dx\,dy \quad (5)$$

$$= \alpha \int_0^2 \left[\frac{1}{3}x^3y + x^2\right]_0^2 dy \quad (6)$$

$$= \alpha \int_0^2 \left[\frac{8}{3}y + 4\right] dy \quad (7)$$

$$= \alpha \left[\frac{4}{3}y^2 + 4y\right]_0^2 \quad (8)$$

$$= 4\alpha \left[\frac{4}{3} + 2\right] = \frac{10}{9} = 1.111 \quad (9)$$

$E[X] = E[Y]$ by symmetry.

$$E[XY] = \int_0^2 \int_0^2 xyp(x,y)\,dx\,dy \quad (10)$$

$$= \alpha \int_0^2 \int_0^2 (x^2y^2 + 2xy)\,dx\,dy \quad (11)$$

$$= \alpha \int_0^2 \left[\frac{1}{3}x^3y^2 + x^2y\right]_0^2 dy \quad (12)$$

$$= \alpha \int_0^2 \left[\frac{8}{3}y^2 + 4y\right] dy \quad (13)$$

$$= \alpha \left[\frac{8}{9}y^3 + 2y^2\right]_0^2 \quad (14)$$

$$= 2\alpha \left[\frac{32}{9} + 4\right] = \frac{34}{27} = 1.259 \quad (15)$$

(cont.

$$
\begin{aligned}
E[X^2] &= \int_0^2 \int_0^2 x^2 p(x,y)\, dx\, dy && (16)\\
&= \alpha \int_0^2 \int_0^2 (x^3 y + 2x^2)\, dx\, dy && (17)\\
&= \alpha \int_0^2 \left[\frac{1}{4}x^4 y + \frac{2}{3}x^3\right]_0^2 dy && (18)\\
&= \alpha \int_0^2 \left[4y + \frac{16}{3}\right] dy && (19)\\
&= \alpha \left[2y^2 + \frac{16}{3}y\right]_0^2 && (20)\\
&= 2\alpha \left[4 + \frac{16}{3}\right] = \frac{14}{9} = 1.556 && (21)
\end{aligned}
$$

(iii) $E[X] = E[Y]$ as the problem is symmetric in x and y.

If X and Y were independent random variables, we would be able to write $E[XY] = E[X]E[Y] = E[Y]^2$. As they are not independent there is no reason for this to be the case.

6 Take the mgf for the binomial distribution and replace p with λ/n:

$$g(z) = [1-p+pz]^n = [1+p(z-1)]^n = \left[1+\frac{\lambda}{n}(z-1)\right]^n$$

Now consider the limit of $\left[1+\frac{u}{n}\right]^n$ as $n \to \infty$:

$$\lim_{n\to\infty}\left[1+\frac{u}{n}\right]^n = \lim_{n\to\infty}\left[1+n\frac{u}{n}+\frac{1}{2!}n(n-1)\left(\frac{u}{n}\right)^2+\frac{1}{3!}n(n-1)(n-2)\left(\frac{u}{n}\right)^3+\dots\right]$$

$$= \lim_{n\to\infty}\left[1+u+\frac{1}{2!}1(1-1/n)u^2+\frac{1}{3!}1(1-1/n)(1-2/n)u^3+\dots\right] = 1+u+\frac{1}{2!}u^2+\frac{1}{3!}u^3+\dots = e^u$$

Thus, we see that

$$\lim_{n\to\infty}\left[1+\frac{\lambda}{n}(z-1)\right]^n = e^{\lambda(z-1)}$$

which is precisely the mgf for a Poisson distribution with parameter λ.

(b) Assume that the probability that any given chocolate is deformed is $p = 0.1$, therefore $q = 1-p = 0.9$. If X is the random variable representing the number of deformed chocolates in a sample of 10, the probability of $X > 2$ is given by a binomial distribution

$$P(X>2) = 1-[P(X=0)+P(X=1)+P(X=2)]$$

giving

$$= 1-\left[\frac{10!}{0!10!}p^0q^{10}+\frac{10!}{1!9!}p^1q^9+\frac{10!}{2!8!}p^2q^8\right] = 1-[0.3487+0.3874+0.1937] = 0.070$$

We can also do this by approximating the distribution by a Poisson distribution with parameter $\lambda = np = 10\times 0.1 = 1$. Thus

$$P(X=r) = \frac{\lambda^r e^{-\lambda}}{r!} \quad\Longrightarrow\quad P(X>2) = 1-\left[\frac{1^0e^{-1}}{0!}+\frac{1^1e^{-1}}{1!}+\frac{1^2e^{-1}}{2!}\right]$$

$$= 1-[0.3679+0.3679+0.184] = 0.080$$

We can therefore see that the results are in reasonable agreement.

(c) Now repeat the calculations of (b) with $n = 5$ and $p = 0.2$ (therefore $q = 0.8$).

$$P(X > 2) = 1 - \left[\frac{5!}{0!5!}p^0q^5 + \frac{5!}{1!4!}p^1q^4 + \frac{5!}{2!3!}p^2q^3\right] = 0.0578$$

In this new case $\lambda = np = 5 \times 0.2 = 1$

thus, $P(X > 2)$ is the same as in the first case: 0.080.

In this case the agreement is no longer reasonable reasonable, and there is a much a greater percentage error.

(d) Recall that the mean and variance for the binomial distribution are np and npq respectively, and for the Poisson distribution are $\lambda = np$ and $\lambda = np$. Thus, unless q is close to 1 (ie p is small), the variances of the distributions will differ. Thus, in order for our Poisson approximation to be a decent one, we will need both a large n and a small p. In general, $p \leq 0.1$ is taken to be a condition which means the approximation is good. In practice, anything above about $n = 8$ means that the distribution is tending towards normal. We see from the above results that indeed the case with a larger value of p and a smaller value of n produced a worse estimate. Better answers might also refer quantitatively to the results in (b) and (c).

END OF PAPER