22B1053 - 22B1205 - 22B1025

Q.7 let X be multinomial distribution random varible, with paramaters $(P_1, P_2, \ldots, P_m, n)$.

let $X = (X_1, X_2, X_3, \ldots, X_m)$

where X_i represent the no. of trials that produced the i^{th} outcome, which are also random varidbles.

let C be covariance matrix of X, where C_{ij} is $Cov(X_i, X_j)$.

now by def,

$$Cov(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)]$$

(where μ_i, μ_j are $E(X_i)$ & $E(X_j)$ respectively)

$$= E[X_iX_j - \mu_i X_j - X_i\mu_j + \mu_i\mu_j]$$

$$= E[X_iX_j] - E[\mu_i X_j] - E[X_i\mu_j] + E[\mu_i\mu_j]$$

$$= E(X_iX_j] - \mu_i E[X_j] - E[X_i]\mu_j + \mu_i\mu_j$$

$$= E[X_iX_j] - \mu_i\mu_j - \mu_i\mu_j + \mu_i\mu_j$$

$$= E[X_iX_j] - E[X_i]E[X_j] \quad — \quad (4)$$

Now to get $E[X_i]$'s and $E[X_iX_j]$'s we use MGF.

First we drive the MGF for X.

$$\phi_X(t) = \phi_X(t_1, t_2, \ldots, t_m)$$

$$= \sum_{k_1+k_2+k_3 \cdots + k_m} \frac{n!}{k_1!k_2!k_3!\cdots k_m!} P_1^{k_1} P_2^{k_2} \cdots P_m^{k_m} e^{t_1k_1} e^{t_2k_2} \cdots e^{t_mk_m}$$

$$= \sum_{k_1+k_2 \cdots + k_m} \frac{n!}{k_1!k_2!\cdots k_m!} (P_1e^{t_1})^{k_1} (P_2e^{t_2})^{k_2} \cdots (P_me^{t_m})^{k_m}$$

now by a multinomial theorem,

$$\phi_X(t) = (P_1e^{t_1} + P_2e^{t_2} \cdots + P_me^{t_m})^n$$

Now as $\phi_X(t) = E(e^{tX}) = E(e^{\Sigma t_iX_i})$

$$\therefore \frac{\partial\phi_X(t)}{\partial t_i} = E[X_i e^{\Sigma t_iX_i}]$$

$$\therefore \left.\frac{\partial\phi_X(t)}{\partial t_i}\right|_{t=0} = E[X_i] \quad — \quad (1)$$

we also have,

for $i \neq j$, $$\frac{\partial^2\phi_X(t)}{\partial t_i\partial t_j} = \frac{\partial(E[X_ie^{\Sigma t_iX_i}])}{\partial t_j} = E[(X_i)(X_j)e^{\Sigma t_iX_i}]$$

$$\therefore \left.\frac{\partial^2\phi_X(t)}{\partial t_i\partial t_j}\right|_{t=0} = E[X_iX_j] \quad — \quad (2)$$

and

$$\frac{\partial \phi_x(t)}{\partial^2 t_i} = \frac{\partial\left(E[x_i e^{\Sigma t_i x_i}]\right)}{\partial t_i} = E\left[x_i^2 e^{\Sigma t_i x_i}\right]$$

$$\left.\frac{\partial^2 \phi_x(t)}{\partial^2 t_i}\right|_{t=0} = E[X_i^2] \qquad — \quad (3)$$

using (1), (2), (3) we find C_{ij} using (4).

for C_{ij} we first find $\frac{\partial^2 \phi_x(t)}{\partial t_i \partial t_j}$, $\frac{\partial \phi_x(t)}{\partial t_i}$

$$\frac{\partial \phi_x(t)}{\partial t_i} = \frac{\partial(\Sigma P_j e^{t_j})^n}{\partial t_i} = n(\Sigma P_j e^{t_j})^{n-1} P_i e^{t_i}$$

for $i \neq j$, $$\frac{\partial^2 \phi_x(t)}{\partial t_i t_j} = \frac{\partial\left(n P_i e^{t_i}(\Sigma P_k e^{t_k})^{n-1}\right)}{\partial t_j} = n P_i e^{t_i}(n-1)(\Sigma P_k e^{t_k})^{n-2} P_j$$

$$\frac{\partial^2 \phi_x(t)}{\partial^2 t_i} = \frac{\partial\left(n P_i e^{t_i}(\Sigma P_k e^{t_k})^{n-1}\right)}{\partial t_i} = n P_i e^{t_i}(\Sigma P_k e^{t_k})^{n-1} + n(n-1)P_i^2 e^{2t_i}(\Sigma P_k e^{t_k})^{n-2}$$

$$\therefore E(X_i) = \left.\frac{\partial \phi_x(t)}{\partial t_i}\right|_{t=0} = n P_i$$

for $i \neq j$, $$E(X_i X_j) = \left.\frac{\partial^2 \phi_x(t)}{\partial t_i \partial t_j}\right|_{t=0} = n(n-1) P_i P_j$$

$$E(X_i^2) = \left.\frac{\partial^2 \phi_x(t)}{\partial^2 t_i}\right|_{t=0} = n P_i + n(n-1) P_i^2$$

$\therefore$ for $i \neq j$, $$C_{ij} = E(X_i X_j) - E(X_i)E(X_j)$$
$$= n(n-1)P_i P_j - n^2 P_i P_j$$
$$C_{ij} = P_i P_j (n^2 - n - n^2) = -n P_i P_j$$

$\because$ C_{ij} is symmetry $(P_i P_j = P_j P_i)$

$\therefore C_{ij} = C_{ji}$.

$$C_{ii} = E(X_i^2) - E^2(X_i)$$
$$= nP_i + (n^2 - n)P_i^2 - n^2 P_i^2$$
$$C_{ii} = n P_i (1 - P_i)$$

$\therefore$ The final covariance matrix $C = \begin{cases} C_{ij} = -nP_i P_j & \text{for } i \neq j \\ C_{ii} = nP_i(1-P_i) \end{cases}$