Joint Distribution of Discrete Random Variables (DRV)

Def Given n DRVs, $\vec{X} = [X_1, X_2, \ldots X_n]$ their joint pmf (jpmf) is given by

$$P_{\vec{X}}(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, X_2 = x_2, \ldots X_n = x_n)$$

Example – $n = 2$ X, Y

$$P_{X,Y}(x, y) = P(X = x, Y = y) \quad (1)$$

Can recover individual distribution (or distribution of a sub collection) using Total Probability Formula

<u>Def</u> (Marginalization) case $n=2$

$$P_X(x) = \sum_y P_{X,Y}(x,y) \qquad (2)$$

$$P_Y(y) = \sum_x P_{X,Y}(x,y) \qquad (3)$$

(Note: $P_X(x) = P(X=x)$

(*) $\underset{\text{TPF}}{=} \sum_y P(X=x, Y=y) \qquad (4)$

$= \sum_y P_{X,Y}(x,y)$)

Can "Marginalize out" any subcollection by summing.

For example if $n=4$

$$P_{X_1,X_3}(x_1,x_3) = \sum_{x_2}\sum_{x_4} P_{X_1,X_2,X_3,X_4}(x_1,x_2,x_3,x_4) \quad (5)$$

<u>Expectation</u> Given

$\mathbf{X} = [X_1, X_2, \ldots, X_n]$ and

$g: \mathbb{R}^n \longrightarrow \mathbb{R}$ let

$Y = g(X_1, X_2, \ldots, X_n)$

Then

$$E[Y] = \sum_{x_1}\cdots\sum_{x_n} g(x_1, x_2, \ldots, x_n)\, P_{\mathbf{X}}(x_1, x_2, \ldots, x_n) \quad (6)$$

Case n = 2 - Given X, Y,

$$g : \mathbb{R}^2 \to \mathbb{R}^1$$

$$Z = g(X, Y)$$

Then

$$E[Z] = \sum_x \sum_y g(x,y) P_{X,Y}(x,y) \quad (7)$$

Example $Z = X + Y$

$$E[Z] = E[X+Y] = \sum_x \sum_y (x+y) P_{X,Y}(x,y)$$

(8)

$$= \sum_x \sum_y x P_{X,Y}(x,y) + \sum_x \sum_y y P_{X,Y}(x,y)$$

(9)

$$= \sum_x x \sum_y P_{X,Y}(x,y) + \sum_y y \sum_x P_{X,Y}(x,y)$$

$$= \sum_x x\, P_X(x) + \sum_y y\, P_Y(y) \qquad (10)$$

$$= E(X) + E(Y). \qquad (11)$$

$$\Rightarrow \quad E[X+Y] = E[X] + E[Y]. \qquad (12)$$

This easily generalizes:

$$(*) \quad E[X_1 + X_2 \dots + X_n] = \sum_{j=1}^{n} E[X_j]. \qquad (13)$$

<u>Note</u> - We already know $E[cX] = cE[X]$

$\Rightarrow$ $E[\cdot]$ is a "Linear Transformation"

$$*) \quad E\left[\sum_{j=1}^{n} c_j X_j\right] = \sum_{j=1}^{n} c_j E[X_j] \qquad (14)$$

What about the Variance of a sum?

Let $Z = X + Y$

$$Var(Z) = E[(Z - \mu_Z)^2] = E[Z^2] - \mu_Z^2 \quad (15)$$

<u>Now</u> $\mu_Z^2 = (\mu_X + \mu_Y)^2$

$$= \mu_X^2 + 2\mu_X\mu_Y + \mu_Y^2 \quad (16)$$

$$E[Z^2] = E[(X+Y)^2]$$

$$= E[X^2 + 2XY + Y^2] \quad (17)$$

$$= E[X^2] + 2E[XY] + E[Y^2]$$

$\Rightarrow$

$$\mathrm{Var}(Z) = E(Z^2) - \mu_Z^2$$

$$= E[X^2] + 2E[XY] + E[Y^2] - (\mu_X^2 + 2\mu_X\mu_Y + \mu_Y^2) \tag{18}$$

$$= E[X^2] - \mu_X^2 + E[Y^2] - \mu_Y^2 + 2(E[XY] - \mu_X\mu_Y) \tag{19}$$

$$= \mathrm{Var}[X] + \mathrm{Var}[Y] + 2(E[XY] - \mu_X\mu_Y)$$

$$\Longrightarrow$$

$$\mathrm{Var}(Z) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2(E[XY] - \mu_X\mu_Y) \tag{20}$$

So **If** $E[XY] = \mu_X \mu_Y$ (21)

(mean of product = product of means)

Then

$$Var[Z] = Var[X] + Var[Y]. \quad (22)$$

<u>Unfortunately</u> This is **not** necessarily so!! –

<u>Def</u> Covariance of X, Y.

$$Cov[X,Y] = E[(X-\mu_X)(Y-\mu_Y)]$$

(23)

Properties of Cov(· , ·)

(1) $\mathrm{Cov}(X,Y) = E[XY] - \mu_X \mu_Y$

*(2) $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$

*(3) Cov is Bilinear

$$\mathrm{Cov}(aX+bY, Z) = a\,\mathrm{Cov}(X,Z) + b\,\mathrm{Cov}(Y,Z)$$

$$\mathrm{Cov}(X, aY+bZ) = a\,\mathrm{Cov}(X,Y) + b\,\mathrm{Cov}(X,Z)$$

*(4.) $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$

(5) $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)$

*(6) $\mathrm{Cov}(X+a, Y) = \mathrm{Cov}(X,Y)$

Defn Covariance Matrix -

Given $\vec{X} = [X_1, X_2, \ldots, X_n]$

define $\Sigma \in \mathbb{R}^{n \times n}$

$$\Sigma_{ij} = \mathrm{Cov}(X_i, X_j) \quad (25)$$

Note $\Sigma_{ii} = \mathrm{Var}[X_i]$

$$\begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & * \\ & & & \\ & * & & \\ & & & \sigma_n^2 \end{bmatrix} \quad (26)$$

!!

Properties of Σ

(1) $\Sigma = \Sigma^T$ (symmetric)

(2) $\Sigma \geq 0$ (positive semidefinite)

$(\vec{a}^T \Sigma \vec{a} \geq 0$ if $\vec{a} \neq \vec{0})$

(3) Given $[a_1, a_2 \ldots a_n]^T = \vec{a} \in \mathbb{R}^n$

(*) $\mathrm{Var}\left(\sum_{i=1}^{n} a_i X_i\right) = \vec{a}^T \Sigma \vec{a}$ (27)

In particular if $\vec{a} = [1, 1, \ldots, 1]^T$

*) $\mathrm{Var}(X_1 + X_2 \ldots X_n) = \sum_i \sum_j \Sigma_{ij}$ (28)

$= \sum_i \mathrm{Var}(X_i) + 2 \sum_{i<j} \mathrm{Cov}(X_i, X_j)$

(4)

Note for $n = 2$

$$\Sigma = \begin{bmatrix} \sigma_X^2 & Cov(X,Y) \\ Cov(X,Y) & \sigma_Y^2 \end{bmatrix} \quad (29)$$

Interpretation of Covariance

$$Cov(X,Y) \begin{cases} > 0 \\ < 0 \end{cases} \text{, if } X \text{ and } Y$$

tend to be on $\begin{cases} \text{same} \\ \text{opposite} \end{cases}$ side

of their means with high probability

Application Hedge Funds like Stocks with **negative** **Covariances**

Proof of Property 1 of Cov.

$$\text{Cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] \quad (30)$$

$$= E[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y] \quad (31)$$

$$= E[XY] - \mu_X E[Y] - \mu_Y E[X] + \mu_X \mu_Y \quad (32)$$

$$= E[XY] - \mu_X \mu_Y - \mu_Y \mu_X + \mu_X \mu_Y \quad (33)$$

$$= E[XY] - \mu_X \mu_Y. \quad (34)$$

Back to $Var(X+Y)$ -

so $Var(X+Y) = Var(X) + Var(Y)$
$+ 2\,Cov(X,Y)$ (35)

Now - if $Cov(X,Y) \left\{ \begin{matrix} > \\ < \end{matrix} \right. 0$

$\left\{ \begin{matrix} \text{more} \\ \text{less} \end{matrix} \right.$ Variation than sum of

Variances! -

Why? if $Cov(X,Y) < 0$ when "one pos other neg" - tends to cancel out variation!

What about if $Cov(X, Y) = 0$

(*) $\Rightarrow Var(X+Y) = Var(X) + Var(Y)$ (36)

Note In this case

$$\sigma_{X+Y} \neq \sigma_X + \sigma_Y \quad (37)$$

but rather

$$\sigma_{X+Y} = \sqrt{\sigma_X^2 + \sigma_Y^2} \quad (38)$$

<u>Defn</u> We say two DRV's are <u>Independent</u> if for any two subsets $A, B \subseteq \mathbb{R}$

$$P(X \in A, Y \in B) = P(X \in A) P(Y \in B) \quad (39)$$

<u>Note</u> If X, Y independent

Then

$$P_{X,Y}(x,y) = P(X = x, Y = y)$$
$$= P(X = x) P(Y = y) \quad (40)$$
$$= P_X(x) P_Y(y).$$

i.e. joint pmf = product of marginal pmf's.

This idea generalises to n random variables $(n > 2.)$ in obvious way.

<u>Example</u> Roll red and green fair dice. R = dots on top of red die G = dots on top of green die. —

<u>Then</u> R, G independent

$$P_{R,G}(i,j) = P(R=i, G=j) = \frac{1}{36} = \frac{1}{6}\cdot\frac{1}{6} \quad (4.1)$$

$$= P(R=i)P(G=j) = P_R(i)P_G(j) \quad (42)$$

<u>But</u> <u>if</u> $X = R+G$, $Y = RG$.

<u>Then</u>

$$P_{X,Y}(3,2) = P(R+G=3,\ RG=2) \quad (43)$$

$$= P(\{R=2, G=1\} \cup \{R=1, G=2\}) \quad (44)$$

$$= P(\{R=2, G=1\}) + P(\{R=1, G=2\}) \quad (45)$$

$$= \frac{1}{36} + \frac{1}{36} = \frac{1}{18} \quad (46)$$

$$\neq \left(\frac{1}{18}\right)^2 = P(R+G=3)P(RG=2) \quad (47)$$

$$= P_X(3)P_Y(2). \Rightarrow \text{dependent} \quad (48)$$

Theorem If X, Y independent.

Then $E(XY) = E(X)E(Y) = \mu_X \mu_Y$

and (1) $Cov(X, Y) = 0$

(2) $Var(X+Y) = Var(X) + Var(Y)$

In general, if X_i, X_j ind. $i \neq j$

(3) $Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$

Pf

$$E(XY) = \sum_x \sum_y xy \, P_{X,Y}(x,y) \quad (49)$$

$$\underset{\text{ind.}}{=} \sum_x \sum_y xy \, P_X(x) P_Y(y)$$

$$= \sum_x x P_X(x) \sum_y y P_Y(y) \quad (50)$$

$$= E(X)E(Y) = \mu_X \mu_Y. \quad (51)$$

Question: We know X, Y ind.

$$\Rightarrow E(XY) = E(X)E(Y). \quad (52)$$

Is it true that $E(XY) = E(X)E(Y)$

$\Rightarrow X, Y$ independent?

Answer No! in general

Example

$$X = \begin{cases} 1 & \frac{1}{3} \\ 0 & \frac{1}{3} \\ -1 & \frac{1}{3} \end{cases} \quad (53)$$

$$(*) \quad Y = |X| = \begin{cases} 1 & \frac{2}{3} \\ 0 & \frac{1}{3} \end{cases} \quad (54)$$

$$(*) \quad XY = X|X| = \begin{cases} 1 & \frac{1}{3} \\ 0 & \frac{1}{3} \\ -1 & \frac{1}{3} \end{cases} \quad (55)$$

$$\Rightarrow \quad E(X) = 0 \qquad E(Y) = \frac{2}{3} \tag{56}$$

$$E(XY) = 0 \tag{57}$$

$$\Rightarrow \quad E(XY) = E(X)\,E(Y). \tag{58}$$

BUT

$$P_{X,Y}(1,1) = P(X=1, Y=1) \tag{59}$$

$$= P(X=1, |X|=1) \tag{60}$$

$$= P(X=1) = \frac{1}{3} \tag{61}$$

$$P_X(1)\,P_Y(1) = P(X=1)\,P(Y=1) \tag{62}$$

$$= \frac{1}{3}\,\frac{2}{3} = \frac{2}{9} \tag{63}$$

$$\Rightarrow \quad P_{X,Y}(1,1) \neq P_X(1)\,P_Y(1) \tag{64}$$

$\Rightarrow$ dependent.

Defn Correlation Coefficient of X, Y

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} \qquad (65)$$

Properties of Correlation Coeff.

(1) X, Y independent $\Rightarrow \rho_{X,Y} = 0$

(2) $\rho_{X,Y} = 0 \Rightarrow X, Y$ uncorrelated
(but not necessary independent)

(3) $-1 \leq \rho_{X,Y} \leq 1$

Pf of (2)

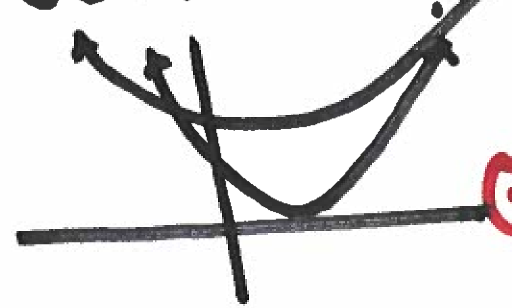
Let α be a scaler

$$0 \le Var(X + \alpha Y) \qquad (66)$$

$$= Var(X) + \alpha^2 Var(Y) \qquad (67)$$

$$+ 2\alpha \, Cov(X, Y) \qquad (68)$$

$$= p(\alpha) - \text{quadratic!} \qquad (69)$$



(*) $\Rightarrow b^2 - 4ac \le 0 \qquad (70)$

$$4\, Cov(X, Y)^2 - 4\, Var(Y)\, Var(X) \le 0 \qquad (71)$$

$$\Rightarrow Cov(X, Y)^2 \le Var(X)\, Var(Y) \qquad (72)$$

$$\Rightarrow |Cov(X, Y)| \le \sigma_X \sigma_Y \qquad (73)$$

Note if $Y = aX + b$

$$\mathrm{Cov}(X, Y) = \mathrm{Cov}(X, aX + b) \quad (74)$$

$$= \mathrm{Cov}(X, aX) + \mathrm{Cov}(X, b) \quad (75)$$

$$= a\,\mathrm{Cov}(X, X) = a\,\sigma_X^2 \quad (76)$$

$$\Rightarrow \rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y} \quad (77)$$

$$= \frac{a\,\sigma_X^2}{\sigma_X |a| \sigma_X} = \frac{a}{|a|} \quad (78)$$

$$= \mathrm{sgn}(a) = \begin{cases} 1 & a > 0 \\ -1 & a < 0 \end{cases} \quad (79)$$

$\Rightarrow$ if X, Y linearly related

$\Rightarrow \rho_{X,Y} = \pm 1$ (80)

if X, Y unrelated -
i.e. independent -

$\rho_{X,Y} = 0$ (81)

$\rho_{X,Y}$ measure of dependence of
X, Y - in some sense

$|\rho_{X,Y}|$ largest (± 1) when
X, Y "very" related - in
most dependent fashion!

Example 1 $X \sim NB(r, p)$

$$\Rightarrow \quad X = X_1 + X_2 \dots + X_r \quad (82)$$

$$X_i \underset{iid}{\sim} G(p) \quad (83)$$

iid - independent identically distributed.

$$\Rightarrow \quad E[X] = E\left[\sum_{i=1}^{r} X_i\right] \quad (84)$$

(*)

$$= \sum_{i=1}^{r} E[X_i] = \sum_{i=1}^{r} \frac{1}{p} = \frac{r}{p}. \quad (85)$$

$$Var(X) = Var\left(\sum_{i=1}^{r} X_i\right) \quad (86)$$

(*)

$$\underset{\text{ind}}{=} \sum_{i=1}^{r} Var(X_i) = \sum_{i=1}^{r} \frac{(1-p)}{p^2} = \frac{r(1-p)}{p^2}. \quad (87)$$

$$\sigma_X = \sqrt{Var(X)} \qquad (88)$$

$$= \frac{\sqrt{r(1-p)}}{p}. \qquad (89)$$

<u>Example 2a</u> See pages 27a 27b

<u>Example 2b</u> $X \sim H(N, n, m)$

<u>Method 1</u> Number distinguished objects (d.o.s) $1, 2, 3, \ldots, m$ so you can tell them apart.

Indicator Random Variables

For $i = 1, 2, \ldots m$, Let

$$I_i = \begin{cases} 1 & i\text{th d.o. in sample} \\ 0 & \text{if not} \end{cases} \qquad (90)$$

Example 2e

$X \sim B(n, p)$ (90)

Define $I_i = \begin{cases} 1 & \text{success on } i \text{th trial} \\ 0 & \text{otherwise} \end{cases}$ (91)

$i = 1, 2, \ldots, n$

(*) Then 1. $I_i \sim B(1, p)$ $\quad i = 1, 2, \ldots, n$

(*) 2. $I_i \sim B(1, p)$ iid.

(*) 3. $X = \sum_{i=1}^{n} I_i$ (92)

(*) 4. $E(X) = E\left[\sum_{i=1}^{n} I_i\right]$ (93)

$= \sum_{i=1}^{n} E[I_i]$ (94)

$= \sum_{i=1}^{n} p = np.$ (95)

(*) 5. $Var(X) = Var\left(\sum_{i=1}^{n} I_i\right) \underset{\text{Ind}}{=} \sum_{i=1}^{n} Var(I_i)$ (96)

$$\ast = \sum_{i=1}^{n} p(1-p) \quad (97)$$

$$\ast = np(1-p). \quad (98)$$

$\#$

(*) Then $$X = \sum_{i=1}^{m} I_i \qquad (100)$$

$$\Rightarrow E(X) = E\left[\sum_{i=1}^{m} I_i\right] \qquad (101)$$

$$= \sum_{i=1}^{m} E(I_i). \qquad (102)$$

Now

$$E[I_i] = 1 \cdot P(i\text{th d.o. in sample}) + 0 \cdot P(i\text{th d.o. not in sample}) \qquad (103)$$

$$= P(i\text{-th d.o. in sample}) \qquad (103a)$$

$$* \quad = \frac{\binom{N-1}{n-1}}{\binom{N}{n}} \qquad (104)$$

← put i-th one in sample Then choose the rest

$$* \quad = \frac{\dfrac{(N-1)!}{(n-1)!\,(N-1-(n-1))!}}{\dfrac{N!}{n!\,(N-n)!}} = \frac{n}{N}. \qquad (105)$$

$$\Rightarrow E(X) = \sum_{i=1}^{m} E[I_i] \qquad (106)$$

$$= \sum_{i=1}^{m} \frac{n}{N} = \frac{mn}{N}. \qquad (107)$$

What about Variance? -

$$Var(X) = Var\left(\sum_{i=1}^{m} I_i\right) \qquad (108)$$

$$= \sum_{i=1}^{m} Var(I_i) + 2\sum_{i<j} Cov(I_i, I_j) \qquad (109)$$

$$Var(I_i) = E[I_i^2] - \left(\frac{n}{N}\right)^2 \qquad (110)$$

★ $E[I_i^2] = 1^2 \cdot P(i\text{-th d.o. in sample})$
$+ 0^2 \cdot P(i\text{-th d.o. not in sample})$

$$= 1 \cdot \frac{n}{N} = \frac{n}{N}. \qquad (111)$$

$$\Rightarrow \operatorname{Var}(I_i) = \frac{n}{N} - \left(\frac{n}{N}\right)^2 \qquad (112)$$

$$= \frac{n}{N}\left(1-\frac{n}{N}\right) = \frac{n}{N}\left(\frac{N-n}{N}\right) \qquad (113)$$

$$= \frac{n(N-n)}{N^2}. \qquad (114)$$

$$\operatorname{Cov}(I_i, I_j) = E[I_i I_j] - E[I_i]E[I_j] \qquad (115)$$

$$= E[I_i I_j] - \left(\frac{n}{N}\right)^2 \qquad (116)$$

$$I_i I_j = \begin{cases} 1 & i\text{th and } j\text{th d.o. in sample} \\ 0 & \text{otherwise} \end{cases} \qquad (117)$$

$$\Rightarrow E[I_i I_j] = P(i\text{-th and } j\text{th d.o. in sample}) \qquad (118)$$

$$= \frac{\binom{N-2}{n-2}}{\binom{N}{n}} \qquad (119)$$

← put i-th, jth d.o. in sample, then choose the rest!

$$\Rightarrow E[I_i I_j] = \frac{\dfrac{(N-2)!}{(n-2)!\,(N-2-(n-2))!}}{\dfrac{N!}{n!\,(N-n)!}} \quad (120)$$

$$= \frac{n(n-1)}{N(N-1)}. \quad (121)$$

$$\Rightarrow \operatorname{Cov}(I_i, I_j) = \frac{n(n-1)}{N(N-1)} - \left(\frac{n}{N}\right)^2 \quad (122)$$

$$\Rightarrow$$

$$\operatorname{Var}(X) = \sum_{i=1}^{m} \operatorname{Var}(I_i) + 2 \sum_{i<j} \operatorname{Cov}(I_i, I_j) \quad (123)$$

$$= \sum_{i=1}^{m} \frac{n(N-n)}{N^2} + 2 \sum_{i<j} \left[\frac{n(n-1)}{N(N-1)} - \left(\frac{n}{N}\right)^2\right] \quad (124)$$

$$= \frac{mn(N-n)}{N^2} + 2\,\frac{m(m-1)}{2}\left(\frac{n(n-1)}{N(N-1)} - \left(\frac{n}{N}\right)^2\right) \quad (125)$$

$$* = n\frac{M}{N}\left(1-\frac{M}{N}\right)\left(\frac{N-n}{N-1}\right). \quad (126)$$

(note there are

$$* \quad 1+2+3\ldots+m-1 = \frac{m(m-1)}{2} \quad (127)$$

terms in $\sum_{i<j}$

Method 2

For $i = 1, 2, \ldots, n$ Let

$$J_i = \begin{cases} 1 & \text{i-th element in Sample is distinguished} \\ 0 & \text{Otherwise} \end{cases} \quad (128)$$

Then if $X \sim H(N, n, m)$ (129)

$$X = \sum_{i=1}^{n} J_i \quad (130)$$

Now find $E[X]$ and $Var[X]$ as in method 1.

Example 3

Note X, Y independent,

$$f: \mathbb{R} \to \mathbb{R}, \quad g: \mathbb{R} \to \mathbb{R} \qquad (131)$$

$\Rightarrow$ $f(X), g(Y)$ independent.

$$P(f(X) \in A, g(Y) \in B)$$
$$= P(X \in f^{-1}(A), Y \in g^{-1}(B)) \qquad (132)$$
$$\underset{\text{ind}}{=} P(X \in f^{-1}(A)) P(Y \in g^{-1}(B)) \qquad (133)$$
$$= P(f(X) \in A) P(g(Y) \in B). \qquad (134)$$

Claim X, Y independent –

Then $M_{X+Y}(t) = M_X(t) M_Y(t)$ (135)

pf

$$M_{X+Y}(t) = E[e^{t(X+Y)}] \quad (136)$$

$$= E[e^{tX+tY}] \quad (137)$$

$$= E[e^{tX}e^{tY}] \quad (138)$$

$$= E[e^{tX}]E[e^{tY}] \quad (139)$$

$$= M_X(t)M_Y(t) \quad (140)$$

If $X_i \underset{iid}{\sim} X$, $i = 1, 2, \ldots, n$ $\Rightarrow M_{\sum_{i=1}^{n} X_i}(t) = M_X(t)^n$. (141)

Example $X \sim G(p) \Rightarrow M_X(t) = \dfrac{pe^t}{1-(1-p)e^t}$ (142)

$\Rightarrow$ if $Y \sim NB(r, p) \Rightarrow$

$$Y = \sum_{i=1}^{r} X_i \quad X_i \underset{iid}{\sim} X. \quad (143)$$

$$\Rightarrow M_Y(t) = \left(\frac{pe^t}{1-(1-p)e^t}\right)^r. \quad (144)$$

Distribution of a sum.

$$\mathbb{X} \sim P_{\mathbb{X}} \quad \mathbb{Y} \sim P_{\mathbb{Y}} \tag{145}$$

$$P_{\mathbb{X}+\mathbb{Y}}(z) = P(\mathbb{X}+\mathbb{Y}=z) \tag{146}$$

$$= \sum_x P(\mathbb{X}=x, \mathbb{Y}=z-x) \tag{147}$$

$$* \quad = \sum_x P(\mathbb{Y}=z-x \mid \mathbb{X}=x) P(\mathbb{X}=x) \tag{148}$$

Now if in addition $\mathbb{X}, \mathbb{Y}$ independent

$$P_{\mathbb{X}+\mathbb{Y}}(z) = P(\mathbb{X}+\mathbb{Y}=z) \tag{149}$$

$$= \sum_x P_{\mathbb{X}}(x) P_{\mathbb{Y}}(z-x) \tag{150}$$

$$= (P_{\mathbb{X}} * P_{\mathbb{Y}})(z) \tag{151}$$

Convolution Product.

Example $X \sim B(n,p)$ $Y \sim B(m,p)$ ind

$$Z = X + Y \Rightarrow \mathcal{R}(Z) = \{0, 1, 2, \ldots, m+n\}$$

For $k = 0, 1, 2, \ldots, m+n$:

k - not n - $X + Y = k \Rightarrow X \leq K$ *

$$P_Z(k) = \sum_{j=0}^{k} P_X(j) P_Y(k-j) \qquad (152)$$

$$= \sum_{j=0}^{k} \binom{n}{j} p^j (1-p)^{n-j} \binom{m}{k-j} p^{k-j} (1-p)^{m-(k-j)} \qquad (153)$$

$$= \sum_{j=0}^{k} \binom{n}{j}\binom{m}{k-j} p^k (1-p)^{m+n-k} \qquad (154)$$

$$= p^k (1-p)^{m+n-k} \sum_{j=0}^{k} \binom{n}{j}\binom{m}{k-j} \qquad (155)$$

$$= \binom{m+n}{k} p^k (1-p)^{m+n-k} \sum_{j=0}^{k} \frac{\binom{n}{j}\binom{m}{k-j}}{\binom{m+n}{k}} \qquad (156)$$

Consider -

$$U \sim H(n+m, k, n) \qquad (157)$$

$$P_U(j) = \frac{\binom{n}{j}\binom{m}{k-j}}{\binom{m+n}{k}} \quad j = 0, 1, 2, \ldots, k \qquad (158)$$

$$\Rightarrow 1 = \sum_{j=0}^{k} P_U(j) = \sum_{j=0}^{k} \frac{\binom{n}{j}\binom{m}{k-j}}{\binom{m+n}{k}} \qquad (159)$$

$\Rightarrow$

$$P_Z(k) = \binom{m+n}{k} p^k (1-p)^{m+n-k} \qquad (160)$$

$$k = 0, 1, 2, \ldots, m+n$$

$$\Rightarrow Z \sim B(m+n, p) \qquad (161)$$

—