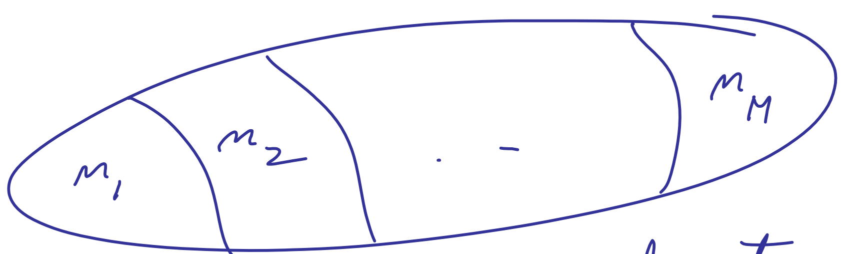
Some Basic Models

# of ways of partitioning $\ni$ m_1 in 1st & m_2 in 2nd

$$= \binom{N}{m_1, m_2} = \frac{N!}{m_1!\, m_2!} \quad \left(= \binom{N}{m_1} \text{ or } \binom{N}{m_2} \right)$$

total # of partitions

$$= \sum_{0 \le m_1 + m_2 = N} \binom{N}{m_1, m_2} = 2^N$$



M subsets

# of ways of partitioning with n_1 in 1st, ..., n_M in the M th is

$$\binom{N}{n_1, \dots, n_M}$$ — multinomial coefficient

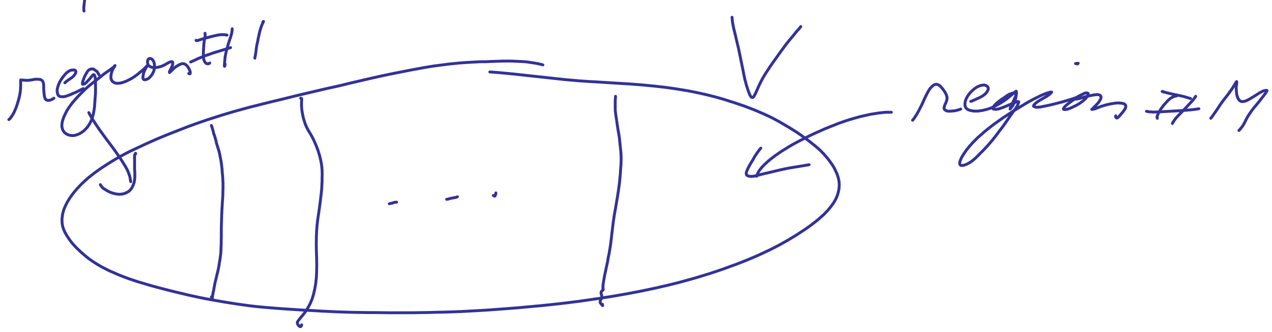
# of partitions

$$= \sum_{0 \le n_1 + \cdots + n_M = N} \binom{N}{n_1, \dots, n_M} = (\overbrace{1 + \cdots + 1}^{M})^N$$

$$= M^N$$

<u>Application</u>

N - particles into M regions



particles placed in V.

# of arrangements of particles $= M^N$

Let $\underset{\sim}{X}$ keep track of which regions
$N \times 1$ the particles are in.

Notice X_i can only take on values $1, \dots, M$. $\longleftarrow$ $\underbrace{\text{ith component of } \underset{\sim}{X}}_{\text{relates to ith particle}}$

There are M^N possible "values" of $\underset{\sim}{X}$.

Assume particles are placed in some random way such that

$$(*) \quad E\big[h(\underset{\sim}{X})\big] = \frac{1}{M^N} \sum_{\substack{\text{all possible} \\ \underset{\sim}{x}}} h(\underset{\sim}{x}), \quad \underline{\underline{\forall h}}$$

real-valued
$(h: \mathbb{R}^N \to \mathbb{R})$

(says particles have no preference in regions)

Special h

$$h(\underset{\sim}{x}) = h_1(x_1)\, h_2(x_2) \cdots h_N(x_N)$$

Plug this into $(*)$ to get

$$E\big[h_1(X_1) \cdots h_N(X_N)\big] = \Big[\frac{1}{M} \sum_{x_1} h_1(x_1)\Big] \cdots \Big[\frac{1}{M} \sum_{x_N} h_N(x_N)\Big]$$

$$= E[h_1(X_1)]\cdots E[h_N(X_N)]$$

$$\text{(x)} \quad E[\underbrace{h_1(X_1)}_{\text{fn of } \underset{\sim}{X}}] = \frac{1}{M^N} \sum_{\text{possible } \underset{\sim}{x}} h_1(x_1) = \frac{1}{M} \sum_{x_1=1}^{M} h_1(x_1)$$

<u>Def'n</u> $Y_1, Y_2, \ldots$ are <u>independent</u> if

$$E[h_1(Y_1)\, h_2(Y_2) \cdots]$$

$$= E[h_1(Y_1)]\, E[h_2(Y_2)] \cdots, \quad \text{“}\forall\text{” } h_i\text{'s}$$

<u>Def'n</u> $A_1, A_2, \ldots$ are independent if $I(A_1), I(A_2), \ldots$ are.

eg Suppose A_1 & A_2 are ind. Then

$$E[I(A_1)\, I(A_2)] = E[I(A_1)]\, E[I(A_2)]$$

$$\Rightarrow \quad P(A_1 A_2) = P(A_1)\, P(A_2)$$

<u>Notice</u> A_1, A_2, A_3 independent

$$\Rightarrow E\left[I(A_1)\, I(A_2)\, \underbrace{h_3}_{\equiv 1}\big(I(A_3)\big) \right]$$

$$= E\big(I(A_1)\, I(A_2) \times 1\big) = P(A_1)\, P(A_2)$$

In fact if $A_1, A_2, \dots$ are ind then

$$P(A_{i_1} A_{i_2} \cdots A_{i_k}) = P(A_{i_1})\, P(A_{i_2}) \cdots P(A_{i_k})$$

(e.g. $1, 2, 3, \dots$

$i_1 < i_2 < i_3 < \dots$

form a subsequence of provided the i's $\in \mathbb{N}$)

Some applications

Call X a <u>counting rv</u> if its range $\subset \{0, 1, 2, \ldots\}$.

$$f(x) = P(X = x) \quad \text{— } \cancel{\text{pf}}$$

Notice $f(x) \geq 0$

$$\sum_x f(x) = 1$$

$$G(z) = E\left(z^X\right) \quad \text{— } \cancel{\text{pgf}}$$

<u>Note</u> — $|G(z)| \leq E\left(|z|^X\right) \leq 1$ if $|z| \leq 1$

— $G(z) = \sum_{k=0}^{\infty} z^k P(X = k)$

— G determines the dist'n of X

eg $X \sim \text{Bernoulli}(p)$, $0 < p < 1$.

$$\left.\begin{array}{l} P(X = 1) = p \\ P(X = 0) = 1 - p \quad (= q) \end{array}\right\} \begin{array}{l} \text{pgf is} \\ G(z) = q + pz \end{array}$$

Let $X_1, X_2, \cdots$ be iid Bernoulli(p) rv's.

$$Y = X_1 + \cdots + X_N \sim \text{binomial}(N, p)$$

$$G_Y(z) = E(z^Y) = E\left(z^{X_1 + \cdots + X_N}\right)$$
$$= E\left(z^{X_1} \cdots z^{X_N}\right)$$
$$= E\left(z^{X_1}\right) \cdots E\left(z^{X_N}\right)$$
$$= (q + pz)^N$$

binomial theorem

$$= \sum_{k=0}^{N} \binom{N}{k} (pz)^k q^{N-k}$$
$$= \sum_{k=0}^{N} \underbrace{\binom{N}{k} p^k q^{N-k}}_{P(Y=k)} z^k$$

pgf of a binomial(N, p) is

$$(q + pz)^N = E(z^Y) = G_Y(z)$$

$$\frac{d}{dz} G_Y(z) \underset{\uparrow}{=} E\left(\frac{d}{dz} z^Y\right) = E\left(Y z^{Y-1}\right)$$

in general
need to justify

$$\frac{d^2}{dz^2} G_Y(z) = E\left(\frac{d^2}{dz^2} z^Y\right) = E\left(Y(Y-1) z^{Y-2}\right)$$

put $z=1$ to get $E(Y)$ + $\underbrace{E[Y(Y-1)]}_{E(Y^2)-E(Y)}$

$$\frac{d}{dz}(q+pz)^N = N(q+pz)^{N-1}p \overset{z=1}{=} Np$$

$$\frac{d^2}{dz^2}(q+pz)^N = N(N-1)(q+pz)^{N-2}p^2 \overset{z=1}{=} N(N-1)p^2$$

$$\therefore \left.\begin{array}{l} E(Y) = Np \\ E(Y^2) - E(Y) = N(N-1)p^2 \end{array}\right\} \Rightarrow E(Y^2) = N(N-1)p^2 + Np$$

$$Var(Y) = E(Y^2) - (E(Y))^2 = N(N-1)p^2 + Np - N^2p^2$$
$$= Npq$$

$$G(z) = P(X=0) + P(X=1)z + \cdots$$

So $G(0) = P(X=0)$

$$G^{(1)}(0) = P(X=1)$$

$$G^{(2)}(0) = 2!\, P(X=2)$$

$$\vdots$$

$$G^{(k)}(0) = k!\, \underline{P(X=k)}$$

$$\sum_{k=0}^{\infty} \overbrace{e^{-\lambda} \frac{\lambda^k}{k!}}^{\text{Poisson}(\lambda) \text{ probabilities}} z^k$$

Back to

$$(q+pz)^N = \left[1 + \frac{\overbrace{Np}^{\lambda}(z-1)}{N}\right]^N \approx \overbrace{e^{\lambda(z-1)}}$$

for large N.

If $X \sim \text{Poisson}(\lambda)$

$$E(X) = \text{Var}(X) = \lambda$$

$$f(x) = e^{-\lambda} \frac{\lambda^x}{x!}, \quad x = 0, 1, \ldots$$

$$= 0, \quad \text{ow}$$

Remark: Finite sum of ind Poisson rv's will be Poisson.

binomial

iid Bernoulli trials yielding

$$\underset{2\times 1}{\underset{\sim}{X}_1}, \underset{2\times 1}{\underset{\sim}{X}_2}, \ldots, \underset{2\times 1}{\underset{\sim}{X}_N} \longrightarrow \text{iid } \underset{\sim}{X}$$

$p_1 + p_2 = 1$

Note possible values of $\underset{\sim}{X}$ are $\overbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}^{p_2}$ or $\overbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}^{p_1}$

~~jt pdf~~ jt pgf of the components of $\underset{\sim}{X}$ or just the the pgf of $\underset{\sim}{X}$

$$G_{\underset{\sim}{X}}(\underset{\sim}{z}) = E\left(\underset{\sim}{z}^{\underset{\sim}{X}}\right) = E\left(z_1^{\text{1st comp of } \underset{\sim}{X}} \, z_2^{\text{2nd comp of } \underset{\sim}{X}}\right)$$

$$= z_2 p_2 + z_1 p_1$$

$$= p_1 z_1 + p_2 z_2$$

= pgf of a "vector Bernoulli"

Now set

$$\underset{\sim}{Y} = \underset{\sim}{X}_1 + \cdots + \underset{\sim}{X}_N$$

$$\Rightarrow \quad G_{\underset{\sim}{Y}}(\underset{\sim}{z}) = \left(p_1 z_1 + p_2 z_2\right)^N = \sum_{y_1 + y_2 = N} \binom{N}{y_1, y_2} p_1^{y_1} p_2^{y_2} z_1^{y_1} z_2^{y_2}$$

$$\Rightarrow \quad P(\underset{\sim}{Y} = \underset{\sim}{y}) = \binom{N}{y_1, y_2} p_1^{y_1} p_2^{y_2}$$

Note $z_2 = 1 \Rightarrow (p_1 z_1 + p_2)^N$ is the pgf of the 1st component of $\underset{\sim}{X}$ (this is the binomial(N, p_1) as you knew it).

Extend to the multinomial

$$\underset{M\times 1}{\underset{\sim}{X}_1}, \underset{M\times 1}{\underset{\sim}{X}_2}, \ldots, \underset{M\times 1}{\underset{\sim}{X}_N} \quad \text{iid } \underset{\sim}{X}$$

$\underset{\sim}{X}$ has one component $= 1$ & the rest 0. The probability that the 1 is in the ith place we call p_i $(i = 1, \ldots, M)$. The pgf of $\underset{\sim}{X}$ is

$$(p_1 z_1 + p_2 z_2 + \cdots + p_M z_M)$$

& letting

$$\underset{\sim}{Y} = \underset{\sim}{X}_1 + \cdots + \underset{\sim}{X}_N$$

we get

$$G_{\underset{\sim}{Y}}(\underset{\sim}{z}) = (p_1 z_1 + \cdots + p_M z_M)^N$$

$$\Rightarrow \quad P(Y = y) = \underbrace{\binom{N}{y}}_{\binom{N}{y_1, \ldots, y_M}} p_1^{y_1} \cdots p_M^{y_M}$$

Note $\frac{\partial^2}{\partial z_1 \partial z_2} G(\underset{\sim}{z}) = E\left(Y_1 z_1^{Y_1 - 1} Y_2 z_2^{Y_2 - 1} z_3^{Y_3} \cdots\right)$ works for all counting rvectors

$\underset{\sim}{z} = \underset{\sim}{1}$: $E(Y_1 Y_2)$

etc....

Can use to get the $\mathrm{cov}(Y_i, Y_j)$ for counting rvec's $\underset{\sim}{Y}$.

<u>Properties of covariances</u> $\left(\mathrm{cov}(X, Y) = E(XY) - E(X)E(Y)\right)$

verify
$$\begin{cases} \mathrm{cov}(X, X) = \mathrm{Var}(X) \\ \mathrm{cov}(X+c, Y+d) = \mathrm{cov}(X, Y) \\ \mathrm{cov}(aX, bY) = ab\,\mathrm{cov}(X, Y) \\ \mathrm{cov}\left(\sum_i X_i, \sum_j Y_j\right) = \sum_{i,j} \mathrm{cov}(X_i, Y_j) \end{cases}$$

eg $X_1 \sim \mathrm{Poisson}(\lambda_1)$, $X_2 \sim \mathrm{Poisson}(\lambda_2)$, $X_3 \sim \mathrm{Poisson}(\lambda_3)$

$$\left.\begin{aligned} U &= X_1 + X_2 \\ V &= X_2 + X_3 \end{aligned}\right\} \text{easy to get the pgf}$$

$$= E\left(z_1^U z_2^V\right)$$

$$= E\left(z_1^{X_1 + X_2} z_2^{X_2 + X_3}\right)$$

$$= E\left(z_1^{X_1} (z_1 z_2)^{X_2} z_2^{X_3}\right)$$

$$= E\left(z_1^{X_1}\right) E\left(z_1 z_2\right)^{X_2} E\left(z_2^{X_3}\right)$$

$$= e^{\lambda_1(z_1 - 1)} e^{\lambda_2(z_1 z_2 - 1)} e^{\lambda_3(z_2 - 1)}$$

– Use this to get $E(UV)$ – imp
& then get $\text{cov}(U, V)$ } 15 minutes

Another way $\text{cov}(U, V) = \text{cov}(X_1 + X_2, X_2 + X_3)$

$$= \text{cov}(X_1, X_2) + \text{cov}(X_1, X_3) + \text{cov}(X_2, X_2) + \text{cov}(X_2, X_3)$$

$$= \text{cov}(X_2, X_2) = \text{Var}(X_2) = \lambda_2$$

Note X, Y ind $\Rightarrow E(XY) = E(X)E(Y) \Rightarrow \text{cov}(X, Y) = 0$

eg $Z \sim N(0,1)$. Let $X = Z$, $Y = Z^2$. Then
X & Y are dependent but

$$E(XY) = E(Z Z^2) = E(Z^3) = 0 = \underbrace{E(X)}_{0} \underbrace{E(Y)}_{1}$$

& so X & Y are uncorrelated.