Lecture 20: Applications of combinatorics to Probability.

Actually, many enumeration problems first arose as probability puzzles, with books written by Montmort and de Moivre, and letters between the Bernoulli's and others. We will, however, do the reverse.

Def (Probability measure) Let Ω be a set and $\mathcal{B} \subseteq 2^{\Omega}$ (a set of subsets of Ω). A probability measure is a function $p : \mathcal{B} \to [0,1]$ s.t.

(1) $p(A) \geq 0$ for all $A \in \mathcal{B}$ ($p : \to [0,1]$ implies this)

(2) For a countable family $(A_1, A_2, \dots)$ in $\mathcal{B}$, pairwise disjoint, $p(\cup A_i) = \sum_{i=1} p(A_i)$

(3) $p(\Omega) = 1$.

- Countable: 1-to-1 correspondence with a subset of $\mathbb{N}$. finite is allowed.
- $A \in \mathcal{B} \to A =$ 'Event'. $p(A) =$ prob of event A.

Remark (Philosophy / Approach to Probability).
Let say $A_i = \{a_i\}$, $\Omega = \{a_1, a_2, \dots\}$

$p(\cup A_i) = \sum p(A_i)$
$\quad \| $
$p(\Omega) = 1$.

Whenever we have a sum $\sum a_n = A$, $a_n \geq 0$. (finite or inf.)
then $\frac{1}{A} \sum a_n = 1$ so we can define a prob measure. – if a_n counts something, automatically $a_n \geq 0$.
Explains title of lecture.

Prop.

1. $p(A) \leq 1$ for all $A \in \mathcal{B}$
2. $p(A \cup A^c) = 1 = p(A) + p(A^c)$
   $\quad$ " Ω
3. $p(\emptyset) = 0$.
4. If $A \subseteq B$ then $p(A) \leq p(B)$
   - $B = A \cup (B \setminus A)$ is a disjoint union.
5. If $A \subset B$, then $p(B - A) = p(B) - p(A)$
6. $p(A \cup B) = p(A) + p(B) - p(A \cap B)$.
7. $p(A_1 \cup A_2 \cup \dots A_n) = \dots$ (using PIE).

Proof. Ex.

Ex 1 (Uniform distribution).
Interpret $\sum_{i \in I_n} 1 = n$ as prob. $\Leftrightarrow \sum_{i \in I_n} \frac{1}{n} = 1$.

$\Omega = \{\omega_1, \dots, \omega_n\}$
$p(\{\omega_i\}) = 1/n$ for $i = 1, 2, \dots n$
$E = \{e_1, e_2, \dots, e_m\} \subset \Omega$
$p(E) = p(\{e_1\}) + \dots + p(\{e_m\}) = m/n$.
"equally likely"

Def (Independent events) Two events are independent if $p(A \cap B) = p(A)p(B)$.

Def (Random Variable) A function $X : \Omega \to \mathbb{R}$ is called a r.v.
Ω: countable $\to$ discrete r.v.

Example: Coin toss $\Omega = \{H, T\}$: $X(H) = 0$, $X(T) = 1$.
Allows quantification. Say you get Rs 1 for H, Rs 10 for T.
$X(H) = 1$, $X(T) = 10$.

Prop: Let X, Y be r.v. Then $X+Y$, $X-Y$, XY, X/Y $(Y \neq 0)$, $aX + bY$ $(a, b \in \mathbb{R})$ are all r.v.

Proof: Algebra of functions

Notation: $P_r(X = x) =$ Prob of $E = \{\omega \in \Omega \mid X(\omega) = x\}$.
Can do $P_r(1 \leq X \leq 5)$ etc.

Remark: ① A "distribution" is a way of assigning probability to all elementary events.

Example (Repeated trials) Sp we toss coins. X_k – r.v. for k^{th} toss
$$X_k = \begin{cases} 1 & \text{if } k^{th} \text{ toss is } H \\ 0 & T. \end{cases}$$

$S_n = X_1 + X_2 + \dots + X_n$ represents the # of heads in multiple tosses of a coin.

Ex. (Geometric sum) Recall: $\sum_{n=0}^{\infty} \frac{1}{2^n} = 1$

Prob interpretation: Toss a coin until you get a heads.
Say $X =$ # of tosses until this happens.
$$p(X = n) : \underbrace{\frac{1}{2} \cdot \frac{1}{2} \cdots \frac{1}{2}}_{n-1 \text{ tails.}} \cdot \underbrace{\frac{1}{2}}_{\text{heads}} = \frac{1}{2^n}$$

More generally: Sp $p(\text{Heads}) = p$ (biased coin).
$p(\text{Tails}) = q$ $(= 1 - p)$.

$p(X = n) = q^{n-1} \cdot p = (1-p)^{n-1} p$.

Note:
$$\sum_{n=0}^{\infty} (1-p)^{n-1} p = p(1 + (1-p) + (1-p)^2 + \dots)$$
$$= \frac{p}{1 - (1-p)} = \frac{p}{p} = 1$$

So it's a legal prob distribution called Geometric distribution.

Def: (Expected value). Let X be a r.v. (discrete). The expectation of X, $E(X) := \sum_{x \in \Omega} x \, P(X = x)$
$=$ weighted average of x, weighted by probability.

Eg: What is avg of: $1, 1, 2, 2, 2, 3$?
Ave $= (2 \cdot 1 + 3 \cdot 2 + 3)/6$

Ex: Find $E(X)$ in above: $\sum_{n=0}^{\infty} n \, q^{n-1} p \overset{\text{Ex.}}{=} \frac{1}{p}$

Hint. Use derivatives.

Ex (Binomial dist / Bernoulli trials). Recall: $\sum_{k=0}^{n} \binom{n}{k} x^k = (1+x)^n$.

Probabilistic Interpretation: n coins are tossed. Prob of heads $= p$
tails $= q$ $(= 1 - p)$

$X_k =$ k^{th} toss is heads
$S_n = X_1 + X_2 + \dots + X_n$
$P(S_n = k) = \binom{n}{k} p^k q^{n-k}$, $0 \leq k \leq n$.
$=$ prob there are k heads exactly.
$\to$ Binary strings with exactly k 1s

Note:
$$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1$$

$$E(S_n) = \sum_{k=0}^{n} k \binom{n}{k} p^k q^{n-k} = \;?$$

Consider $\sum_{k=0}^{n} \binom{n}{k} x^k = (1+x)^n$

Apply $D = x\frac{d}{dx}$ $\Rightarrow \sum_{k=0}^{n} k \binom{n}{k} x^k = x \cdot n \cdot (1+x)^{n-1}$

$$E(S_n) = q^n \sum_{k=0}^{n} k \binom{n}{k} (p/q)^k = q^n \cdot \frac{p}{q} (n) \cdot (1 + p/q)^{n-1}$$
$$= np. \quad \square$$

Moral: $E(X) \leftarrow$ use $D = x\frac{d}{dx}$ on "probability g.f."
$P(x) = \sum_{n=0}^{\infty} P_n x^n$. $\therefore D.P(x) = \sum_{n=0} n P_n x^n$.

Ex: (Birthday Paradox). Prob of 2 people having same birthday. Put your birthday on a piece of paper and lets see if you get same birthday.
$n = 23$, prob. $> 1/2$.

Ex. Birthday paradox: n people
No birthday in common: $S(n) = \dfrac{365(365-1)\dots(365-n+1)}{(365)^n}$

$1 - S(n)$: $n = 23$ > 0.5
$n = 40$: prob 0.89