Lecture 15

Method of Indicators

Binomial mean (again)

Tail sum formula

Boole's inequality

Markov's inequality

Addition rule

$$E\{X+Y\} = E\{X\}+E\{Y\}$$

- Makes sense.
- Requires no conditions beyond existence of the expectations.
- Needs a proof.
- Immediate extension:

$$E\{X_1 + X_2 + \ldots + X_n\} \;=\; E\{X_1\} + E\{X_2\} + \ldots + E\{X_n\}.$$

Indicator functions

$$1_A = \begin{cases} 1 & \text{if } A \text{ occurs} \iff \omega \in A \\ 0 & \text{otherwise} \iff \omega \notin A \end{cases}$$

$$E\{1_A\} = P\{A\}$$

Expected number of events

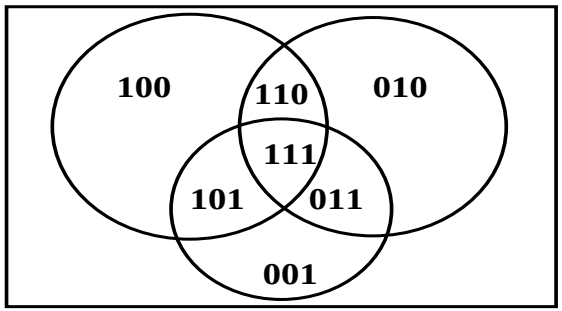
- If $A_1, A_2, \ldots$ are any events, then

$$E\{\text{Number of events occuring}\} = E\left\{\sum_k 1_{A_k}\right\} = \sum_k P\{A_k\}$$

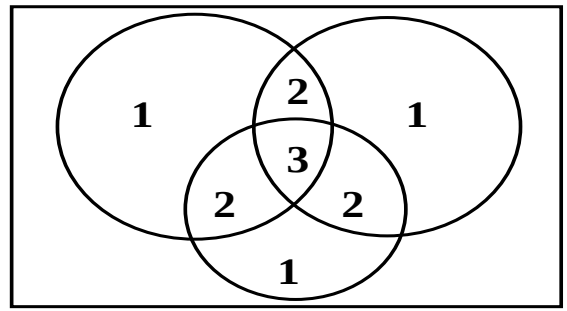
- Example: A system has 3 components. Find the expected number working at any time.

  - Write 1_k for the indicator of the event that compent number k working.

  - Then $X = 1_1 + 1_2 + 1_3$ is the <u>total</u> number working.



Values of $(1_1, 1_2 1_3)$



Values of $X = 1_1 + 1_2 + 1_3$

$$E\{X\} = \sum_{k=1}^{3} P\{\text{Component number } k \text{ is working}\}.$$

The mean of a binomial

- We seek $E\{X\}$, where $X \sim \text{Bin}(n, p)$.
- $1_k,\ k = 1, \ldots, n$ are indicators of "success" on k-th trial.
- $X = \sum_1^n 1_k$.
- $E\{1_k\} = p$.

$$E\{X\} \;=\; E\left\{\sum_1^n 1_k\right\} \;=\; \sum_1^n E\{1_k\} \;=\; np$$

Sum of binomials

- $X \sim \text{Bin}(n_1, p_1)$ and $Y \sim \text{Bin}(n_2, p_2)$
  $\Rightarrow\ E\{X + Y\} = n_1 p_1 + n_2 p_2$.
- $X + Y$ does not have a binomial distribution, <u>unless</u> $p_1 = p_2 = p$, and then $X + Y \sim \text{Bin}(n_1 + n_2, p)$.

Tail sum formula

- If X takes values $\{0, 1, 2, \ldots, n\}$ ($n = \infty$ is OK) then

$$\boxed{E\{X\} \;=\; \sum_{k=1}^{n} P\{X \geq k\}}$$

- Example: $X \sim \text{Geom}(p), \quad P\{X = k\} = pq^{k-1}, \quad k \geq 1..$

$$P\{X \geq k\} \;=\; \sum_{j=k}^{\infty} pq^{j-1} \;=\; pq^{k-1} \sum_{j=0}^{\infty} q^{j} \;=\; \frac{pq^{k-1}}{1-q} \;=\; q^{k-1}$$

$$E\{X\} \;=\; \sum_{k=1}^{\infty} P\{X \geq k\} \;=\; \sum_{k=1}^{\infty} q^{k-1} \;=\; \sum_{k=0}^{\infty} q^{k} \;=\; \frac{1}{1-q} \;=\; \frac{1}{p}.$$

Mean of a negative binomial

- T_r = number of trials until r-th success in Bernoulli trials.
- $p_T(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \qquad k \geq r.$
- $E\{T_r\} = \sum_{k=r}^{\infty} k \, \binom{k-1}{r-1} p^r (1-p)^{k-r}, \qquad k \geq r.$

- W_j = time between $(j-1)$-st and j-th successes.
- $T_r = W_1 + W_2 + \ldots + W_r.$
- $E\{W_j\} = 1/p.$
- $\Rightarrow \quad E\{T_r\} = r/p.$

$$\boxed{E\{X\} \;=\; \sum_{k=1}^{n} P\{X \geq k\}}$$

Proof 1 (Algebra)

$$\begin{aligned}
E\{X\} &= \sum_{k=0}^{n} k p_k \\
&= \quad p_1 \\
&\qquad +p_2 \quad +p_2 \\
&\qquad +p_3 \quad +p_3 \quad +p_3 \\
&\qquad \ldots \quad \ldots \quad \ldots \quad \ldots \\
&\qquad +p_n \quad +p_n \quad +p_n \quad + \ldots \quad +p_n \\
&= P\{X \geq 1\} + P\{X \geq 2\} + \ldots + P\{X \geq n\} \\
&= \sum_{k=1}^{n} P\{X \geq k\}
\end{aligned}$$

$$\boxed{E\{X\} \;=\; \sum_{k=1}^{n} P\{X \geq k\}}$$

Proof 2 (Indicators)

- Set $A_j = \{X \geq j\}, \quad j \geq 1.$
- $\{X = k\} \iff A_j$ occurs $\forall\ 1 \leq j \leq k$,
  <u>and</u> A_j does not occur $\forall\ k < j \leq n$.
- Thus $\{X = k\} \iff$ number of A_j ocurring is <u>exactly</u> k.

$$\begin{aligned} E\{X\} &= E\{\text{Number of } A_j \text{ ocurring}\} \\ &= \sum_{j=1}^{n} P\{A_j\} \\ &= \sum_{j=1}^{n} P\{X \geq j\} \end{aligned}$$

Expectation of a minimum

- Four dice are rolled, and M is the <u>minimum</u> of the four numbers. Find $E\{M\}$.
- $E\{M\} \;=\; \sum_{k=1}^{6} P\{M \geq k\}$.
- Let X_k be the number on the k-th die.

$$\begin{aligned} P\{M \geq k\} &= P\{X_1 \geq k,\ X_2 \geq k,\ X_3 \geq k,\ X_4 \geq k\} \\ &= \left(\frac{6-k+1}{6}\right)^4 \end{aligned}$$

$$\begin{aligned} E\{M\} &= P\{M \geq 1\} + P\{M \geq 2\} + \ldots + P\{M \geq 6\} \\ &= \left(\frac{6}{6}\right)^4 + \left(\frac{5}{6}\right)^4 + \left(\frac{4}{6}\right)^4 + \left(\frac{3}{6}\right)^4 + \left(\frac{2}{6}\right)^4 + \left(\frac{1}{6}\right)^4 \approx 1.76 \end{aligned}$$

Alternative computation

- $E\{M\} = \sum_{k=1}^{6} kP\{M = k\}$.

$$
\begin{aligned}
P\{M = k\} &= P\{M \geq k\} - P\{M \geq k+1\} \\
&= \left(\frac{6-k+1}{6}\right)^4 - \left(\frac{6-k}{6}\right)^4
\end{aligned}
$$

BUT

$$
\begin{aligned}
\sum_{k=1}^{6} kP\{M = k\} &= \sum_{k=1}^{6} k\left[P\{M \geq k\} - P\{M \geq k+1\}\right] \\
= \; & 1\left[P\{M \geq 1\} - P\{M \geq 2\}\right] + 2\left[P\{M \geq 2\} - P\{M \geq 3\}\right] + \\
& 3\left[P\{M \geq 3\} - P\{M \geq 4\}\right] + \ldots + 6\left[P\{M \geq 6\} - P\{M \geq 7\}\right] \\
= \; & \sum_{k=1}^{6} P\{M \geq k\}
\end{aligned}
$$

$$\boxed{\text{Boole's inequality: } P\{X \geq 1\} \leq E\{X\}.}$$

(If X takes values in $\{0, 1, 2, \ldots\}$.)

$$P\{X \geq 1\} \;=\; \sum_{k \geq 1} P\{X = k\} \;\leq\; \sum_{k \geq 0} k P\{X = k\} \;=\; E\{X\}$$

$$\boxed{\text{Markov's inequality: } P\{X \geq a\} \leq E\{X\}/a.}$$

(If X takes only <u>non-negative</u> values.)

$$a P\{X \geq a\} \;=\; \sum_{x_k \geq a} a P\{X = x_k\} \;\leq\; \sum_{x_k \geq 0} x_k P\{X = x_k\} \;=\; E\{X\}$$

Two examples

$$X \sim \mathrm{Geom}(p)$$

$$P\{X = k\} = pq^{k-1}, \quad E\{X\} = 1/p, \quad P\{X \geq k\} = q^{k-1}.$$

$$p \;=\; q \;=\; 0.5$$

k	1	10	100	1,000
$P\{X \geq k\}$	1	2.0×10^{-3}	1.6×10^{-30}	2.8×10^{-300}
Boole ($E\{X\}$)	2.0	–	–	–
Markov ($E\{X\}/k$)	2.0	.2	.02	.002

$$X \sim \mathrm{Pois}(\lambda)$$

$$P\{X = k\} = e^{-\lambda}\lambda^k/k!, \quad E\{X\} = \lambda, \quad P\{X \geq k\} = e^{-\lambda}\sum_{j \geq k} \lambda^j/j!$$

$$\lambda \;=\; 2$$

k	1	10	100	1,000
$P\{X \geq k\}$	$1 - e^{-2} = 0.86$	$\sim 10^{-4}$	$\sim 10^{-130}$	$\sim 10^{-3000}$
Boole	2.0	–	–	–
Markov	2.0	0.2	0.02	0.002

Calculations via *Stirling's formula:* $n! \;\sim\; \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$.