Named Discrete Distribution.

Def'n. Bernoulli trials have two possible outcomes.
They are also known as indicators (or indicator r.v.s).
Usually, the outcomes are encoded by

$$\begin{cases} 1 & \text{for "success"} \\ 0 & \text{for "failure"} \end{cases}$$

Example. Y_i ... result of a throw of a die for $i = 1, 2$.

$$S_{Y_1} = S_{Y_2} = \{1, 2, 3, 4, 5, 6\}$$

Define $W = Y_1 + Y_2$

$$\boxed{I = \begin{cases} 1 & \text{if } W \geq 9 \\ 0 & \text{if } W < 9 \end{cases}}$$

Example.
- Quality control, say, whether a lightbulb is deficient or not
- Indicator of whether a deductible is met

Example. Bernoulli distribution.

y	0	1
$p_Y(y)$	$1-p$	p

$Y \sim b(p)$ w/ $p \in (0,1)$

Example Say that we repeat independently the Bernoulli trial w/ the same p a fixed number of times n.
Then, we count the total number of successes Y.
This distribution is the Binomial distribution.

$$\boxed{Y \sim b(n,p)}$$

for $k \in \{0, \dots, n\}$: $p_k = \mathbb{P}[Y = k] = \binom{n}{k} p^k (1-p)^{n-k}$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

M378K Introduction to Mathematical Statistics
Problem Set #1
Named discrete random variables.

Problem 1.1. Source: Sample P exam, Problem #125.
An electronic system contains three cooling components that operate independently. The probability of each component's failure is 0.05. *The system will overheat if and only if at least two components fail. Calculate the probability that the system will overheat.*

→:

$N \sim b(3, 0.05)$

$$\mathbb{P}[N \geq 2] = \mathbb{P}[N=2] + \mathbb{P}[N=3]$$
$$= \binom{3}{2}(0.05)^2(0.95) + \binom{3}{3}(0.05)^3$$
$$= 3(0.05)^2(0.95) + (0.05)^3$$
$$= \cdots = 0.00725 \quad \square$$

Example. Say we repeat independent Bernoulli trials w/ the same success probability p until the first success.
The random variable Y will denote the total number of failures until the first success.

We write $Y \sim g(p)$

and we call Y geometric.

Set $\boxed{q = 1-p}$

y	0	1	2	....	k	......
$p_Y(y)$	p	$q \cdot p$	$q^2 p$	...	$q^k \cdot p$	....

Problem 1.2. Source: Sample P exam, Problem #462.
Each person in a large population independently has probability p of testing positive for diabetes where $0 < p < 1$. People are tested for diabetes, one person at a time, until a test is positive. Individual tests are independent. Determine the probability that m or fewer people are tested, given that n or fewer people are tested, where $1 \leq m \leq n$.

$\rightarrow$: N'... total # of people tested

↳ SHIFTED geometric r.v. w/ parameter p

i.e., $\boxed{N = N' - 1 \sim g(p)}$ ✓

$$\mathbb{P}[N' \leq m \mid N' \leq n] = \mathbb{P}[N+1 \leq m \mid N+1 \leq n]$$
$$= \mathbb{P}[N \leq m-1 \mid N \leq n-1]$$
$$= \frac{\mathbb{P}[N \leq m-1, N \leq n-1]}{\mathbb{P}[N \leq n-1]}$$
$$= \frac{\mathbb{P}[N \leq m-1]}{\mathbb{P}[N \leq n-1]}$$
$$= \frac{1 - \boxed{\mathbb{P}[N > m-1]}}{1 - \mathbb{P}[N > n-1]}$$
$$= \frac{1 - q^m}{1 - q^n}$$

□

__Example.__ The Poisson Distribution is $\mathbb{N}_0$-valued and has the

pmf $$p_k = e^{-\lambda} \cdot \frac{\lambda^k}{k!} \quad \text{for } k = 0, 1, \dots$$

where λ is a positive parameter