- Lecture#7
- -
- - In this lecture we start considering
- some special probability distributions
- associated to discrete random variables.

Remark: Given a DRV X, then for any function $h: \mathbb{R} \to \mathbb{R}$, $h(X)$ is a DRV, and we can compute the mean and variance:

$$E(h(X)) = \sum_x h(x) f(x)$$

$$V(h(X)) = \sum_x (h(x) - E(h(x)))^2 f(x)$$

where $f(x)$ is the probability mass function of X.

Notable case: $h(x) = ax + b$. Then

$$E(aX+b) = aE(X) + b$$

$$V(aX+b) = a^2 V(X).$$

not true in general that $E(h(X)) = h(E(X))$

Geg: $h(x) = x^2$.

Binomial Distribution

- suppose we have some number n, of trials, i.e. an experiment repeated n times.
- for each trial, we assume two possible outcomes
  - i.e. Heads-Tails, success-Failure, 0-1, Yes-No, win-lose, etc.
- the outcomes are mutually exclusive
  - ex: flipping a coin, you can't have both heads <u>and</u> tails.
- suppose our outcomes are success-failure (the names are irrelevant).
- we assume that each trial is independent and that we are successful with some fixed probability p.
  - it follows that each trial fails with probability $1-p$.
  - Such trials are called "Benoulli Trials"
- <u>set up</u>: n Bernoulli trials with probability of success p.

- We say that X is a binomial random variable with parameters $0 < p < 1$, and $n = 1, 2, 3 \ldots$ (sometimes write X ~~as~~ as $Bin(p, n)$) if
$X =$ "number of successes" in n Bernoulli trials with probability p.

- The probability mass function of X (i.e. of $Bin(p,n)$) is

$$\boxed{f(x) = \binom{n}{x} p^{x} (1-p)^{n-x}}, \quad x = 1, 2, \ldots, n$$

- <u>Interpretation</u>: $f(m) = \binom{n}{m} p^{m} (1-p)^{n-m}$ is the probability of being successful m times in n trials.

- Flipping a single coin (H = success, T = failure) is an example of a Bernoulli trial with $p = 0.5$.

- $Bin(0.5, n) \rightsquigarrow f(x) = \binom{n}{x} \frac{1}{2}^{x} \frac{1}{2}^{n-x}$

  ↳ ex: if we flip a coin twice (two trials) we expect the probability of getting heads twice to be $\left(\frac{1}{2}\right)^2 = \frac{1}{4}$.

  ↳ $f(2) = \binom{2}{2} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^{2-2} = 1 \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^0 = \frac{1}{4}$ ✓

Let's compute the mean and variance of $Bin(p,n)$.

$$\mu = E(Bin(p,n)) \overset{\text{by defn}}{=\!=} \sum_{i=0}^{n} i\, f(i)$$

Sum starts from $i=1$ since the first term with $i=0$ is $0 \cdot f(0) = 0$.

$$= \sum_{i=1}^{n} i \binom{n}{i} p^i (1-p)^{n-i}$$

$$= \sum_{i=1}^{n} \frac{\cancel{i}\, n!}{(n-i)!\, \cancel{i!}_{(i-1)!}} p^i (1-p)^{n-i}$$

$$= \sum_{i=1}^{n} \frac{n!}{(n-i)!\,(i-1)!} p^i (1-p)^{n-i}$$

re index with $j = i-1$ ($i = j+1$)

$$= \sum_{j=0}^{n-1} \frac{n!}{(n-(j+1))!\, j!} p^{j+1} (1-p)^{n-(j+1)}$$

$$= np \left(\sum_{j=0}^{n-1} \frac{(n-1)!}{((n-1)-j)!\, j!} p^j (1-p)^{(n-1)-j} \right)$$

$$= np \left[\sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{(n-1)-j} \right]$$

$$= \underbrace{np\,(p + (1-p))^{n-1}}_{\text{By binomial theorem}} = \boxed{np}$$

For the variance:

$$V(Bin(p,n)) = \sum_{i=1}^{n} (i - np)^2 \binom{n}{i} p^i (1-p)^{n-i}$$

$\vdots$ (exercise)

$$= np(1-p).$$

Ex: Your test on Sep. 30 will consist of 27 multiple choice questions each with 4 choices. Suppose all questions are independent and you need at least 14 correct answers to pass. What is the probability of passing if you guess randomly on every question?

Soln Each question is a Bernoulli trial with $p = 1/4$ (25% chance of success)

- P.M.F is $f(x) = \binom{27}{x}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{27-x}$

- Recall that the cumulative distribution function is

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i).$$

- we want $P(X \geq 14) = 1 - P(X \leq 13)$

- So, it follows that

$$P(X \leq 13) = \sum_{i=1}^{13} f(i)$$

$$= \sum_{i=1}^{13} \binom{27}{i} \left(\frac{1}{4}\right)^{i} \left(\frac{3}{4}\right)^{27-i}$$

$$= \binom{27}{1}\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^{26} + \binom{27}{2}\left(\frac{1}{4}\right)^{2}\left(\frac{3}{4}\right)^{25} + \dots + \binom{27}{13}\left(\frac{1}{4}\right)^{13}\left(\frac{3}{4}\right)^{14}$$

$$\approx 0.997126$$

So $1 - 0.997126 \approx 0.00287\ldots$

So you have a 0.28% chance of passing. <u>**So study!!**</u>

If everyone is guessing the average/mean should be $\underset{n \cdot p}{\frac{27}{4}} \approx 6.75$ correct answers.