3Y03 - Probability and Statistics for Engineering

WS19 Lecture 9

Last time

Mean & Variance of a Discrete r.v.

"where centre is" "how spread out distribution is"

↳ ways to summarise a prob. distribution of a r.v.

Mean $\mu = E(X) = \sum_x x f(x)$ $\qquad \sigma^2 = V(X) = E((X-\mu)^2) = \sum_x (x-\mu)^2 f(x) = \left(\sum_x x^2 f(x)\right) - \mu^2$

$E(\quad) =$ " expected value of "

We can find the expected value of any function $h(X)$ of a discrete r.v. X with p.m.f. $f(x)$:

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

Notice our alt. formula for $V(X) = \sum_x x^2 f(x) - \mu^2$

$$= E(X^2) - (E(X))^2$$

Notice $h(X) = Y$ is itself a (discrete) random variable & has its own p.m.f. $f_h(y)$

& $E(Y) = \sum_y y f_h(y)$. It can be shown that these give same answer.

For example : $E(aX + b) = aE(X) + b$

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

Special Distributions

Discrete

complementary

- Binomial
- Geometric
- Negative Binomial
- Hypergeometric
- Poisson

Binomial Distribution

- Some # of independent "trials"

n

(e.g. single experiment that can be repeated)

- each trial has 2 mutually exclusive possible outcomes (1 or 0, success or failure, H or T, yes or no, win or loss, ...)
- probability of a particular outcome (success)

(is fixed ← p

↑ called a <u>Bernoulli trial</u>

(So we have n Bernoulli trials with prob. of success $= p$.)

- X = # successes
  - called Binomial random variable with parameters n and p

$X \sim \text{Bin}(n, p)$

- X has p.m.f. $f(x) = \underbrace{\binom{n}{x}}_{} p^x (1-p)^{n-x}$

for each $x = 0, \dots, n$

Binomial coefficients
= total # of sequences of length n with x successes, $n-x$ failures

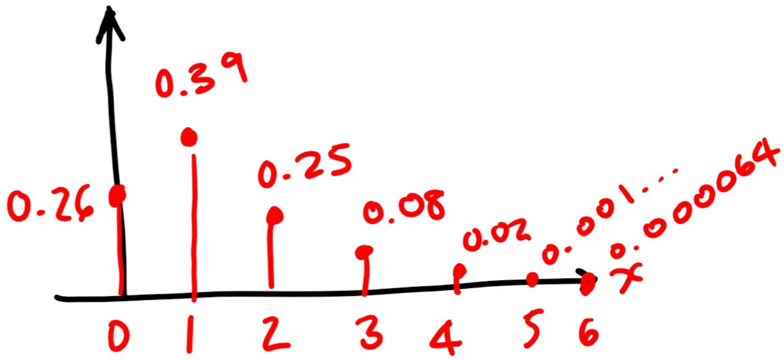
<u>Example</u> Multiple Choice Test with $n = 6$ unrelated questions ; 5 poss. answers to each.
↑ independent

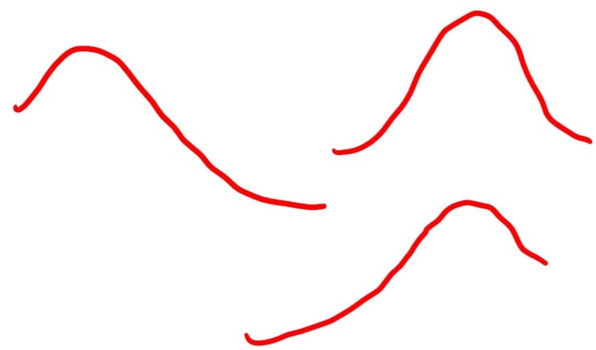
Student randomly guesses answers, answers equally likely so $p = 0.2$

So X = total # of correct answers.

$$P(X=0) = f(0) = \binom{6}{0}(0.2)^{0}(0.8)^{6-0}$$
$$= 0.26$$

$$P(X=4) = f(4) = \binom{6}{4}(0.2)^{4}(0.8)^{6-4}$$
$$= 0.02$$





Notice that the cdf $F(x_0) = P(X \le x_0)$

$$= \sum_{x \le x_0} f(x)$$

$$= \sum_{x \le x_0} \binom{n}{x} p^{x} (1-p)^{n-x}$$

<u>Example</u>

In the Multiple-Choice Example above find $P(X \le 3)$, $P(2 \le X \le 5)$, $P(X \ge 4)$

<u>Solution</u> $P(X \le 3) = \sum_{x \le 3} f(x) = f(0) + f(1) + f(2) + f(3)$

$$= 0.26 + 0.39 + 0.25 + 0.08$$
$$= 0.98.$$

$$P(2 \leq X \leq 5) = \sum_{x \leq 5} f(x) - \sum_{x \leq 1} f(x)$$

" $F(5)$ " $F(1)$

$$= \sum_{x=2}^{5} f(x) = f(2) + f(3) + f(4) + f(5)$$

$$= 0.35.$$

$$P(X \geq 4) = \sum_{x \geq 4} f(x) = f(4) + f(5) + f(6)$$

"

$$1 - P(X \leq 3) = 1 - \sum_{x \leq 3} f(x) = 1 - 0.98 = 0.02.$$

Mean & Variance of Binomial Distribution

If $X \sim \text{Bin}(n, p)$ then notice

$$X = X_1 + X_2 + \ldots + X_n \quad \text{where}$$

$$X_i = \begin{cases} 1 & \text{if } i\text{th trial is success} \\ 0 & \text{if } i\text{th trial is failure} \end{cases}$$

$$\mu = E(X) = E(X_1 + \ldots + X_n)$$

justification later (5-4)

$$= E(X_1) + E(X_2) + \ldots + E(X_n)$$

the p.m.f. of $X_i \sim \text{Bin}(1, p)$

Each $X_i \sim \text{Bin}(1, p)$ so $f_i(x) = \binom{1}{x} p^x (1-p)^{1-x}$

& $E(X_i) = \sum_{x=0}^{1} x f_i(x) = 0 f_i(0) + 1.f_i(1)$

$= 1.\binom{1}{1} p^1 (1-p)^0$

$= p$

So $\underline{\underline{E(X) = np}}$

And $V(X) = V(X_1) + \ldots + V(X_n)$ (justification later!)

$= n \sum_{x=0}^{1} (x-p)^2 \binom{1}{x} p^x (1-p)^{1-x}$

$\vdots$

$= \underline{\underline{np(1-p)}}.$

$= n\left(p^2(1-p) + (1-p)^2 p\right)$

$= n(\cancel{p^2} - \cancel{p^3} + p - \cancel{2}p^2 + \cancel{p^3})$

$= n(p - p^2)$