Lecture 5

Variance
Binomial Experiment
Binominal Random Variable

Index

- Variance
- Bernoulli Random Variable
- Binomial Experiment
- Binomial Random Variable
- Binomial Tables
- Questions

Variance

If the rv X has a set of possible values D and pmf $p(x)$, then the expected value of any function $h(X)$, denoted by $E[h(X)]$ or $\mu_{h(X)}$, is computed by

$$E[h(X)] = \sum_{D} h(x) \cdot p(x)$$

Let X have pmf $p(x)$ and expected value μ. Then the **variance** of X, denoted by $V(X)$ or σ_X^2, or just σ^2, is

$$V(X) = \sum_{D} (x - \mu)^2 \cdot p(x) = E[(X - \mu)^2]$$

The **standard deviation** (SD) of X is

$$\sigma_X = \sqrt{\sigma_X^2}$$

Expectation and Variance

E(X+Y) = E(X) + E(Y) where X and Y are random variables

E(aX) = aE(X) where X is a random variable and a is a constant

E(a) = a where a is a constant

Var(aX) = a² Var(X) where X is a random variable and a is a constant

Var(a) = 0 where a is a constant

$$E(aX + b) = a \cdot E(X) + b$$

Expectation and Variance

$$E(aX + b) = a \cdot E(X) + b$$

$Var(X) = E[(X-u)^2]$
$= E[X^2 + u^2 - 2Xu]$
$= E[X^2] + E[u^2] - 2E[Xu]$
$= E[X^2] + u^2 - 2uE[X]$
$= E[X^2] + u^2 - 2uu$
$= E[X^2] - u^2$

Variance

$$V(aX + b) = \sigma^2_{aX+b} = a^2 \cdot \sigma^2_X \quad \text{and} \quad \sigma_{aX+b} = |a| \cdot \sigma_x$$

In particular,

$$\sigma_{aX} = |a| \cdot \sigma_X, \quad \sigma_{X+b} = \sigma_X$$

Question

Let X be a Bernoulli rv with pmf. Calculate the following

1) $E(X)$
2) $E(X^2)$
3) $V(X)$
4) $E(X^{79})$

Note: Assume $p(1) = p$ and $p(0) = 1-p$

Question

X is a Bernoulli rv with pmf $p(1) = p$ and $p(0) = 1-p$.

1) $E(X) = 1.p + 0.(1-p) = p$
2) $E(X^2) = 1^2.p + 0^2.(1-p) = p$
3) $V(X) = E(X^2) - (E(X))^2 = p-p^2 = p(1-p)$
4) $E(X^{79}) = 1^{79}.p + 0^{79}.(1-p) = p$

Question

Let X = the outcome when fair dice is rolled once. If before the die is rolled you are offered either (1/3.5) dollars or h(X) = 1/X dollars, would you accept the guaranteed amount or would you gamble?

Question

Let X = the outcome when fair dice is rolled once. If before the die is rolled you are offered either (1/3.5) dollars or h(X) = 1/X dollars, would you accept the guaranteed amount or would you gamble?

$E(h(X)) = \frac{1}{6}(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6})$

$E(h(X)) = 1/(2.44)$

So E(h(X)) greater than 1/3.5 so would gamble.

Binomial Experiment

The experiment consists of a sequence of n smaller experiments called trials, where n is fixed in advance of the experiment.

Each trial can result in one of the same two possible outcomes, which we generically denote by success (S) and failure (F).

The trials are independent, so that the outcome on any particular trial does not influence the outcome on any other trial.

The probability of success P(S) is constant from trial to trial; we denote this probability by p.

Binomial Random Variable

The **binomial random variable** $\boldsymbol{X}$ associated with a binomial experiment consisting of n trials is defined as

$$X = \text{ the number of S's among the } n \text{ trials}$$

Because the pmf of a binomial rv X depends on the two parameters n and p, we denote the pmf by $b(x; n, p)$.

Binomial Distribution

$n = 6$ i.e. 6 trials $\qquad P(S) = p$

We want 4 success $\qquad P(F) = 1-p$

$$\begin{array}{cccccc} \underline{S} & \underline{S} & \underline{S} & \underline{S} & \underline{F} & \underline{F} \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ p & p & p & p & 1-p & 1-p \end{array}$$

We want SSSSFF i.e. 1st trial Success, 2nd trial Success, 3rd trial Success, 4th trial Success, 5th trial Failure, 6th Failure to happen simultaneously & each of them is independent trial

$$P(SSSSFF) = P(S)P(S)P(S)P(S)P(F)P(F)$$
$$= p.p.p.p.(1-p).(1-p)$$
$$= p^4(1-p)^2$$

Now SSSSFF is one possibility

No. of ways to choose 4 S out of $6 = {}^6C_4$

So Total probability of 4 S out of 6 trials

$$= {}^6C_4\, p^4 (1-p)^2$$

n=4

Outcome	x	Probability	Outcome	x	Probability
$SSSS$	4	p^4	$FSSS$	3	$p^3(1-p)$
$SSSF$	3	$p^3(1-p)$	$FSSF$	2	$p^2(1-p)^2$
$SSFS$	3	$p^3(1-p)$	$FSFS$	2	$p^2(1-p)^2$
$SSFF$	2	$p^2(1-p)^2$	$FSFF$	1	$p(1-p)^3$
$SFSS$	3	$p^3(1-p)$	$FFSS$	2	$p^2(1-p)^2$
$SFSF$	2	$p^2(1-p)^2$	$FFSF$	1	$p(1-p)^3$
$SFFS$	2	$p^2(1-p)^2$	$FFFS$	1	$p(1-p)^3$
$SFFF$	1	$p(1-p)^3$	$FFFF$	0	$(1-p)^4$

Formulae

$$b(x; n, p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & x = 0, 1, 2, \ldots, n \\ 0 & \text{otherwise} \end{cases}$$

For $X \sim \text{Bin}(n, p)$, the cdf will be denoted by

$$B(x; n, p) = P(X \leq x) = \sum_{y=0}^{x} b(y; n, p) \quad x = 0, 1, \ldots, n$$

If $X \sim \text{Bin}(n, p)$, then $E(X) = np$, $V(X) = np(1-p) = npq$, and $\sigma_X = \sqrt{npq}$ (where $q = 1 - p$).

Binomial Tables

Even for a relatively small value of n, the computation of cumulative binomial probabilities can be tedious.

Hence we use binomial tables.

Binomial Tables tell the cdf.

Question

Suppose that 20% of all copies of a particular textbook fail a certain binding strength test. Let X denote the number among 15 randomly selected copies that fail the test. Then X has a binomial distribution with n=15 and p =0.2.

1. Calculate the probability that at most 8 fail the test is

$$P(X \leq 8) = \sum_{y=0}^{8} b(y; 15, .2) = B(8; 15, .2)$$

c. $n = 15$

		p														
		0.01	**0.05**	**0.10**	**0.20**	**0.25**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.75**	**0.80**	**0.90**	**0.95**	**0.99**
	0	.860	.463	.206	.035	.013	.005	.000	.000	.000	.000	.000	.000	.000	.000	.000
	1	.990	.829	.549	.167	.080	.035	.005	.000	.000	.000	.000	.000	.000	.000	.000
	2	1.000	.964	.816	.398	.236	.127	.027	.004	.000	.000	.000	.000	.000	.000	.000
	3	1.000	.995	.944	.648	.461	.297	.091	.018	.002	.000	.000	.000	.000	.000	.000
	4	1.000	.999	.987	.836	.686	.515	.217	.059	.009	.001	.000	.000	.000	.000	.000
	5	1.000	1.000	.998	.939	.852	.722	.403	.151	.034	.004	.001	.000	.000	.000	.000
	6	1.000	1.000	1.000	.982	.943	.869	.610	.304	.095	.015	.004	.001	.000	.000	.000
x	7	1.000	1.000	1.000	.996	.983	.950	.787	.500	.213	.050	.017	.004	.000	.000	.000
	8	1.000	1.000	1.000	.999	.996	.985	.905	.696	.390	.131	.057	.018	.000	.000	.000
	9	1.000	1.000	1.000	1.000	.999	.996	.966	.849	.597	.278	.148	.061	.002	.000	.000
	10	1.000	1.000	1.000	1.000	1.000	.999	.991	.941	.783	.485	.314	.164	.013	.001	.000
	11	1.000	1.000	1.000	1.000	1.000	1.000	.998	.982	.909	.703	.539	.352	.056	.005	.000
	12	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.996	.973	.873	.764	.602	.184	.036	.000
	13	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.995	.965	.920	.833	.451	.171	.010
	14	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.995	.987	.965	.794	.537	.140

Questions

2. Calculate the probability that exactly 8 fails

$$P(X = 8) = P(X \leq 8) - P(X \leq 7) = B(8; 15, .2) - B(7; 15, .2)$$

Questions

2. Calculate the probability that exactly 8 fails

$$P(X = 8) = P(X \leq 8) - P(X \leq 7) = B(8; 15, .2) - B(7; 15, .2)$$

The result is 0.999 - 0.996 = 0.003.

Question

3. Calculate the probability that at least 8 fail.
4. Calculate the probability that between 4 and 7 fail (4 and 7 inclusive)

Refer the table and calculate the answers

$$
\begin{aligned}
P(X \geq 8) &= 1 - P(X \leq 7) = 1 - B(7; 15, .2) \\
&= 1 - \begin{pmatrix} \text{entry in } x = 7 \\ \text{row of } p = .2 \text{ column} \end{pmatrix} \\
&= 1 - .996 = .004
\end{aligned}
$$

$$
\begin{aligned}
P(4 \leq X \leq 7) &= P(X = 4, 5, 6, \text{ or } 7) = P(X \leq 7) - P(X \leq 3) \\
&= B(7; 15, .2) - B(3; 15, .2) = .996 - .648 = .348
\end{aligned}
$$

Just for Explanation

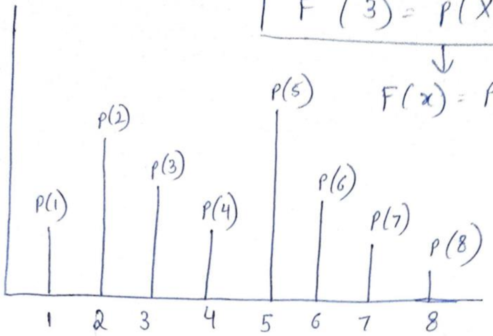
$F(1) = P(X \le 1)$

$F(3) = P(X \le 3)$

$$\downarrow$$

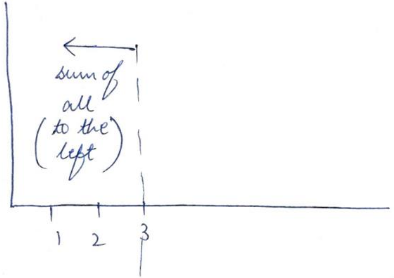
$F(x) = P(X \le x)$

$n = 8$



$F(3) = p(1) + p(2) + p(3) = P(X \le 3)$

X takes value 1, 2 & 3



i) $P(X \le 8) = F(8)$

ii) $P(X = 8) = p(8)$

$F(8) = p(8) + p(7) + p(6) + p(5) + p(4) + p(3) + p(2) + p(1)$

$F(7) = p(7) + p(6) + p(5) + p(4) + p(3) + p(2) + p(1)$

desirable values of $X = 4, 5, 6, 7$

so $P(4 \le X \le 7) = p(4) + p(5) + p(6) + p(7)$

$= F(7) - ?$

$= F(7) - F(3)$

References

Probability and statistics for engineers RA Johnson, I Miller, JE Freund - 2000 - 117.239.47.98

Statistics for business & economics DR Anderson, DJ Sweeney, TA Williams, JD Camm

Probability and statistics for engineering and science J Deovre