Computing formula.

Let Y be a r.v. with mean $E(Y) = \mu$.

An alternative computing formula for the variance is

$$\operatorname{var}(Y) = E\left[(Y-\mu)^2\right]$$
$$= E(Y^2) - \left[E(Y)\right]^2$$

proof

$$E(Y^2 + \mu^2 - 2Y\mu)$$
$$= E(\mu^2) + E(Y^2) - (2Y \cdot \mu)$$
$$= E(\mu^2) + E(Y^2) - \cancel{2E}\; 2\mu E(Y)$$
$$= \mu^2 + E(Y^2) - 2\mu^2$$
$$= E(Y^2) - \mu^2 = E(Y^2) - \left[E(Y)\right]^2$$

Example 3.2.

Compute the ~~pmf~~ variance of Y [Y is the # of gallons of a toxic chemical that is produced per hour. Compute variance of Y

$$\sigma^2 = var(Y) = E(Y^2) - [E(Y)]^2$$

$$= 3.3 - (1.5)^2 = 1.05.$$

The unit is squared gallons. ($gallon^2$)

The standard deviation of Y is

$$\sigma = \sqrt{\sigma^2} = \sqrt{1.05} = 1.25 \text{ (gallon.)}$$

3.2 Binomial distribution.

Bernoulli trials

Many experiments can be envisioned as consisting of a sequence of "trials", where

1, each trial results in a "success" or a "failure"

2, the trials are independent

3. the probability of success, denoted by p, $0<p<1$, is the same on every trial.

e.g

When the circuit boards used in the manufacture of Blue Ray players are tested, the long run percentage of defective boards is 5%

$\underbrace{\text{long run percentage}}_{\text{suggesting probability here.}}$

"trial" – circuit board

"success" – ~~defe~~ defective board is ~~detected~~ detected

"① p = P("success") = P(defective board)

= 0.05

T②

Binomial Distribution

→ A series of Bernoulli trials

→ Count the number of successes as Y.

→ Y is a Random variable

→ Then Y has a distribution

→ That distribution is Binomial Distribution

Suppose that n Bernoulli trials are performed, define

Y = # of successes (out of n trials performed)

We say Y has a binomial distribution with number of trials n and success probability p.

$$Y \sim b(n, p)$$

pmf if $Y \sim b(n,p)$, the pmf of Y is

given by

$$p_Y(y) = \begin{cases} \binom{n}{y} p^y (1-p)^y & y = 0, 1, \ldots, n \\ 0. & \text{o.w} \end{cases}$$

otherwise.

Mean / variance.

If $Y \sim b(n,p)$

then $E(Y) = np$, & $Var(Y) = np(1-p)$

E.g. 3.3. In an agricultural study,

it is determined that 40% of all plots

Respond to a certain treatment. Four plots are observed.

- plot of land $\to$ trial.
- plot respond to treatment $\to$ success
- $p = P(\text{"success"}) = 0.4$
  $= P(\text{"responds to treatment"})$

If the Bernoulli trial assumptions hold (independent plots. Same response probability for each plot..

then Y = the number of plots which respond $\sim b(n, p)$
$\sim b(4, 0.4)$

(a) what is the probability that exactly 2 plots respond ?

$$P(Y=2) = p_Y(2) = \binom{4}{2}(0.4)^2(1-0.4)^{4-2}$$

$$= 6\times(0.4)^2(0.6)^2 = 0.3456$$

(b) What is the probability that at least one responds

$$P(Y\geq 1) = P(Y=1) + P(Y=2) + P(Y=3) + P(Y=4)$$

or $P(Y \geq 1) = 1 - P(Y=0)$

$$= 1 - \binom{4}{0}(0.4)^0(1-0.4)^{4-0}$$

$$= 1 - 1\times 1 \times (0.6)^4 = 0.8704$$

(c) What ~~is~~ are $E(Y)$ and $Var(Y)$

$$E(Y) = np = 1.6$$

$$Var(Y) = 0.96$$
$$= np(1-p)$$

→ You can do it the old way

$$P(Y=0) \rightarrow \binom{4}{0} p^0 (1-p)^4 \rightarrow a$$
$$P(Y=1) \rightarrow \binom{4}{1} p^1 (1-p)^3 \rightarrow b$$
$$P(Y=2) \rightarrow \binom{4}{2} p^2 (1-p)^2 \rightarrow c$$
$$P(Y=3) \rightarrow \binom{4}{3} p^3 (1-p)^1 \rightarrow d$$
$$P(Y=4) \rightarrow \binom{4}{4} p^4 (1-p)^0 \rightarrow e$$

$$E(Y) = P(Y=0)a + P(Y=1)b + \cdots + P(Y=4)e$$
$$= np.$$