Chapter 4. Probability Distribution (B)

February 7, 2011

1 Expectation

Example The male adults in the Village (a very special case).

weight x_k	75.6	75.7	75.8	75.9	76.0	76.1	76.2	76.3	76.4
$P(X = x_k)$	0.075	0.100	0.100	0.175	0.225	0.1375	0.100	0.0625	0.025

Population mean μ

$$\begin{aligned} \mu &= \sum_{x_k} p_k x_k = 75.6 \times 0.075 + 75.7 \times 0.100 + ... + 76.4 \times 0.025 \\ &= \sum_{x_k} P(X = x_k) x_k \ (= 75.965) \end{aligned}$$

Example Suppose 10% of the human population carries the green-eye allele. If we choose 1000 people randomly and let the r.v. X be the number of green-eyed people in the sample. Then the distribution of X is binomial distribution with n $=$ 1000 and p $=$ 0.1 (denoted as $X \sim B(1000, 0.1)$)

In a sample of 1000 people, how many of them are we expecting to have this allele? Clearly the count of individuals that carry the green-eye allele will vary between different samples of 1000 subjects. How much dispersion between the samples can we expect, in terms of the number of individuals carrying this allele? These questions will be answered by computing the mean and the variance (or standard deviation) for this process.

Definition The Expected Value, Expectation or Mean of a discrete random variable X is defined by

$$\mathbf{E}[X] \ \text{(or } E(X), EX) = \sum_{all\ different\ x_i} x_i P(X = x_i).$$

provided that this sum converges absolutely.

Another definition[1] *is*

$$\mathbf{E}[X] = \sum_{\omega \in \Omega} P(\{\omega\}) X(\omega)$$

[1]You can skip this definition

Example Suppose we toss an even die and the upper face is recorded X. We have $p_k = P(X = k) = 1/6$, and

$$EX = 1 * \frac{1}{6} + 2 * \frac{1}{6} + ... + 6 * \frac{1}{6} = 3.5$$

Example If X is a randomly selected value from data $\{x_1, ..., x_N\}$. Then in this special case

$$\mathbf{E}[X] = \mu_x = \sum_{i=1}^{N} x_i/N$$

or (if the data is grouped)

$$\mathbf{E}[X] = \mu_x = \sum_{i=1}^{k} x_i p_i$$

Example [Poisson distribution] If $P(X = r) = e^{-\lambda}\frac{\lambda^r}{r!}$, then $\mathbf{E}[X] = \lambda$.

Because

$$\mathbf{E}[X] = \sum_{r=0}^{\infty} r \cdot e^{-\lambda}\frac{\lambda^r}{r!} = \lambda e^{-\lambda}\sum_{r=1}^{\infty}\frac{\lambda^{r-1}}{(r-1)!} = \lambda e^{-\lambda}e^{\lambda} = \lambda.$$

Example [Binomial distribution] If $P(X = r) = C_n^r p^r(1-p)^{n-r}$ then $\mathbf{E}[X] = np$.

Because

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{r=0}^{n} r \cdot C_n^r p^r (1-p)^{n-r} \\
&= \sum_{r=0}^{n} r \frac{n!}{r!(n-r)!} p^r (1-p)^{n-r} \\
&= n \sum_{r=1}^{n} \frac{(n-1)!}{(r-1)!(n-r)!} p^r (1-p)^{n-r} \\
&= np \sum_{r=1}^{n} \frac{(n-1)!}{(r-1)!(n-r)!} p^{r-1} (1-p)^{n-r} \\
&= np \sum_{r=0}^{n-1} \frac{(n-1)!}{(r)!(n-1-r)!} p^r (1-p)^{n-1-r} \\
&= np \sum_{r=0}^{n-1} C_{n-1}^r p^r (1-p)^{n-1-r} \\
&= np.
\end{aligned}$$

Example Suppose 10% of the human population carries the green-eye allele. If we choose 1000 people randomly and let the r.v. X be the number of green-eyed people in the sample. Then the distribution of X is binomial distribution with $n = 1000$ and $p = 0.1$. In a sample of 1000, how many of them are we expecting to have this allele?

Note that $\mathbf{E}[X]$ is a constant.

Theorem

1. If $X \geq 0$ then $\mathbf{E}[X] \geq 0$.

2. If $X \geq 0$ and $\mathbf{E}[X] = 0$ then $P(X = 0) = 1$.

3. If a and b are constants then $\mathbf{E}[a + bX] = a + b\mathbf{E}[X]$.

4. For any random variables X, Y then $\mathbf{E}[X + Y] = \mathbf{E}[X] + \mathbf{E}[Y]$.

5. $\mathbf{E}[X]$ is the constant which minimises $\mathbf{E}\left[(X - c)^2\right]$.

6. for any function f, $\mathbf{E}[f(X)] = \sum_{x_i} f(x_i)P(X = x_i)$

Proof:

1. Because $X \geq 0$*, so*

$$\mathbf{E}[X] = \sum_x P(X = x)x \geq 0$$

2. If $\exists x$ *with* $P(X = x) > 0$ *and* $x > 0$ *then* $\mathbf{E}[X] \geq xP(X = x) > 0$*,*

therefore $P(X = 0) = 1$*.*

3.

$$\begin{aligned}
\mathbf{E}[a + bX] &= \sum_x (a + bx)\, P(X = x) \\
&= a \sum_x P(X = x) + b \sum_x xP(X = x) \\
&= a + \mathbf{E}[X]\,.
\end{aligned}$$

4. Trivial.

5. Now

$$\begin{aligned}
&\mathbf{E}\big[(X-c)^2\big] \\
&= \mathbf{E}\big[(X-\mathbf{E}[X]+\mathbf{E}[X]-c)^2\big] \\
&= \mathbf{E}\big[[(X-\mathbf{E}[X])^2\big] + 2(X-\mathbf{E}[X])(\mathbf{E}[X]-c) + [(\mathbf{E}[X]-c)]^2] \\
&= \mathbf{E}\big[(X-\mathbf{E}[X])^2\big] + 2(\mathbf{E}[X]-c)\mathbf{E}[(X-\mathbf{E}[X])] + (\mathbf{E}[X]-c)^2 \\
&= \mathbf{E}\big[(X-\mathbf{E}[X])^2\big] + (\mathbf{E}[X]-c)^2.
\end{aligned}$$

This is clearly minimized when $c = \mathbf{E}[X]$.

6. (omitted).

Theorem For any random variables $X_1, X_2, \cdots, X_n$

$$\mathbf{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbf{E}[X_i]$$

Proof:

$$\begin{aligned} \mathbf{E}\left[\sum_{i=1}^{n} X_i\right] &= \mathbf{E}\left[\sum_{i=1}^{n-1} X_i + X_n\right] \\ &= \mathbf{E}\left[\sum_{i=1}^{n-1} X_i\right] + \mathbf{E}[X_n] \end{aligned}$$

Result follows by induction.

Example In the village example, suppose we are going to draw 10 people with replacement. Denote them by $X_1, ..., X_{10}$. We have

$$\mathbf{E}[X_k] = \mu = 75.965$$

Consider

$$\bar{X} = (X_1 + ... + X_{10})/10$$

We still have

$$\mathbf{E}\left[\bar{X}\right] = (\mathbf{E}[X_1] + ... + \mathbf{E}[X_{10}])/10 = \mu = 75.965(kg)$$

What does this calculation imply??

2 Variance

Definition The variance of a random variable X is the number

$$\begin{aligned} \mathbf{var}[X] &= \mathbf{E}\big[(X - \mathbf{E}[X])^2\big] \quad \text{(denoted by } \sigma^2 \text{ or } V[X]), \\ &= \sum_{x_i} (x_i - \mu)^2 p_i \\ \text{Standard Deviation} &= \sqrt{\mathbf{var}[X]} \end{aligned}$$

Example The male adults in the Village (a very special case). X is drawn randomly from the population. $EX = 75.965$kg

$$\mathbf{var}[X] = (75.6 - 75.965)^2 \times 0.075 + ... + (76.4 - 75.965)^2 \times 0.025 = \sigma^2$$

Theorem [Properties of Variance]

1. $\mathbf{var}[X] \geq 0$.

2. If $\mathbf{var}[X] = 0$, then $P(X = \mathbf{E}[X]) = 1$

3. If a, b constants, $\mathbf{var}[(a + bX)] = b^2\mathbf{var}[X]$

   Proof:

$$\begin{aligned}
\mathbf{var}[a + bX] &= \mathbf{E}[a + bX - a - b\mathbf{E}[X]]^2 \\
&= b^2\mathbf{E}[X - \mathbf{E}[X]]^2 \\
&= b^2\mathbf{var}[X]
\end{aligned}$$

4. $\mathbf{var}[X] = \mathbf{E}\left[X^2\right] - \mathbf{E}[X]^2$

Proof:

$$
\begin{aligned}
\mathbf{E}[X - \mathbf{E}[X]]^2 &= \mathbf{E}\left[X^2 - 2X\mathbf{E}[X] + (\mathbf{E}[X])^2\right] \\
&= \mathbf{E}\left[X^2\right] - 2\mathbf{E}[X]\,\mathbf{E}[X] + \mathbf{E}[X]^2 \\
&= \mathbf{E}\left[X^2\right] - (\mathbf{E}[X])^2
\end{aligned}
$$

recall

$$\sum_{i=1}^{N}(x_i - \mu)^2/N = \sum_{i=1}^{N} x_i^2/N - \mu^2$$

Example Let X have the geometric distribution $P(X = r) = pq^r$ with $r = 0, 1, 2, \cdots$ and $p + q = 1$. Then $\mathbf{E}[X] = \frac{q}{p}$ and $\mathbf{var}[X] = \frac{q}{p^2}$.

Proof:

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{r=0}^{\infty} rpq^r = pq \sum_{r=0}^{\infty} rq^{r-1} \\
&= pq \sum_{r=0}^{\infty} \frac{d}{dq}(q^r) = pq \frac{d}{dq}\Big(\frac{1}{1-q}\Big) \\
&= pq(1-q)^{-2} = \frac{q}{p}
\end{aligned}$$

$$
\begin{aligned}
\mathbf{E}\big[X^2\big] &= \sum_{r=0}^{\infty} r^2 p q^r \\
&= pq\left(\sum_{r=1}^{\infty} r(r+1)q^{r-1} - \sum_{r=1}^{\infty} r q^{r-1}\right) \\
&= pq\left(\frac{2}{(1-q)^3} - \frac{1}{(1-q)^2}\right) = \frac{2q}{p^2} - \frac{q}{p}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{var}[X] &= \mathbf{E}\big[X^2\big] - \mathbf{E}[X]^2 \\
&= \frac{2q}{p^2} - \frac{q}{p} - \frac{q^2}{p^2} \\
&= \frac{q}{p^2}
\end{aligned}
$$

Example If $X \sim B(n, p)$, then

$$\mathbf{E}[X] = np, \quad \mathbf{var}[X] = np(1-p)$$

If $X \sim Poi(\lambda)$, then

$$\mathbf{E}[X] = \lambda, \quad \mathbf{var}[X] = \lambda$$

Population and random Variable X

For a population with mean μ and variance σ^2, X is an random sample/observation from the population. Then

$$\mathbf{E}[X] = \mu, \qquad \mathbf{var}[X] = \sigma^2$$

and

$$\text{the distribution of } X = \text{the distribution of the population}$$

Thus to investigate the population, we turn to investigate r.v. X. We also call X the population! $\mathbf{E}[X]$ and $\mathbf{var}[X]$ can be interpreted exactly the same as μ and σ^2 respectively.

3 More Examples

Suppose we are offered to play a game of chance under these conditions: it costs us to play \$1.5 and the awarded prices are \$1, \$2, \$3. Assume the probabilities of winning each price are {0.6, 0.3, 0.1}, respectively. Should we play the game? What are our chances of winning/loosing? Let's let X=awarded price. Then X={1, 2, 3}.

x	1	2	3
P(X=x)	0.6	0.3	0.1

$$\mathbf{E}[X] = 1 * 0.6 + 2 * 0.3 + 3 * 0.1 = 1.5\$$$

The expected return of this game is \$1.5, which equals the entry fee, and hence the game is fair - neither the player nor the house has an advantage in this game (on the long run). Though any n games will produce different outcomes and may give small advantage to one side. However, on the long run, no one will make money.

The variance for this game is computed by

$$\begin{aligned}\mathbf{var}[X] &= (x_1 - 1.5)^2 P(X = x_1) \\ &\quad +(x_2 - 1.5)^2 P(X = x_2) \\ &\quad +(x_3 - 1.5)^2 P(X = x_3) \\ &= 0.45.\end{aligned}$$

Thus, the standard deviation is 0.6708$.

Example Suppose we conduct an (unethical!) experiment involving young couple planning to have children. Suppose, the couples are interested in the number of girls they will have, and each couple agrees to have children until one of the following 2 stopping criteria is met: (1) the couple has at least one child of each gender, or (2) the couple has at most 3 children! Let's denote the r.v. X =number of Girls, and Y =number of kids. Observable Outcomes {BBB, BG; GB; BBG, GGB, GGG}

x	0	1	2	3
P(X=x)	1/8	5/8	1/8	1/8

y	2	3
P(Y=y)	1/2	1/2

$$\mathbf{E}[X] = 0 + 5/8 + 2/8 + 3/8 = 1.25.$$

and

$$\mathbf{E}[Y] = 2.5.$$

Baby gender control paradox: what is the interpretation of this expectations? no matter what strategy is adopted (abortion is prohibited!), the overall ratio of boys/girls does not change! Consider a strategy for a couple: keep giving birth until the first girl.

Can you calculate the variance and standard deviation for this random variable?