Expectation

STAT 211 - 509

2018/09/13

Expectation of discrete random variables

- The expected value is a one-number summary of the distribution of a random variable. Think of it as the average, $\frac{1}{n}\sum_{i=1}^{n} X_i$ of a large number of IID draws $X_1, X_2, \ldots, X_n$.
- The **expected value** (or **mean**) of a discrete random variable X with pmf f is

$$\mathbb{E}[X] = \sum_x x f(x)$$

  Depending on the situation, we may use the Greek symbol μ_x (or just μ, if the context is clear) in place of $\mathbb{E}[X]$.

Example

Let $X \sim Bernoulli(p)$, with pmf

$$f(x) = p^x(1-p)^{1-x}, \quad x \in \{0,1\}, \quad p \in [0,1]$$

Thus X takes the value 1 with probability p and 0 with probability $1-p$. We have

$$\mathbb{E}[X] = \sum_{x=0}^{1} x f(x) = (0 \times (1-p)) + (1 \times p) = p$$

Properties of Expectations

- Let $Y = g(X)$, for some function g. Then

$$\mathbb{E}[Y] = \mathbb{E}[g(X)] = \sum_x g(x) f(x)$$

- If $X_1, X_2, \ldots, X_n$ are random variables and $a_1, a_2, \ldots a_n$ are constants, then

$$\mathbb{E}\left[\sum_i a_i X_i\right] = \sum_i a_i \mathbb{E}[X_i]$$

- For some constant a,

$$\mathbb{E}[a] = a$$

Example

Let $X \sim Binomial(n, p)$. What is $\mathbb{E}[X]$? Applying the expectation definition directly is challenging, because of the complicated form of the pmf. Instead, note that we can write

$$X = \sum_{i=1}^{n} X_i$$

where X_i equals 1 if the ith trial is a success and 0 if it is a failure. Then $\mathbb{E}[X_i] = p$ and

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_i X_i\right] = \sum_i \mathbb{E}[X_i] = np$$

Variance

- Another important way to characterize a distribution is in terms of “spread.” The most common measure of spread is called the variance.

- Let be a random variable with mean . The **variance** of X is

$$Var(X) = \mathbb{E}(X - \mathbb{E}[X])^2 = \sum_x (x - \mathbb{E}[X])^2 f(x)$$

- The **standard deviation** is

$$\sqrt{Var(X)}$$

  Depending on the situation, we may use the Greek symbol σ_X^2, or just σ^2, if the context is clear, in place of $Var(X)$. Similarly, σ_X or σ may be used to indicate the standard deviation.

Properties of Variance

- $$Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$
- If a and b are constants, then
$$Var(aX + b) = a^2 Var(X)$$
- If $X_1, X_2, \ldots, X_n$ are independent and $a_1, a_2, \ldots a_n$ are constants, then
$$Var\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 Var(X_i)$$

Example

Let $X \sim Binomial(n, p)$. What is $Var(X)$? Again write

$$X = \sum_{i=1}^{n} X_i$$

where X_i equals 1 if the nth trial is a success and 0 if it is a failure. Then $E[X_i] = p$ and

$$E[X_i^2] = (p \times 1^2) + ((1-p) \times 0^2) = p$$

So

$$Var(X_i) = E[X_i^2] - p^2 = p(1-p)$$

and

$$Var(X) = Var\left(\sum_i X_i\right) = \sum_i Var(X_i) = np(1-p)$$

Sample Mean and Variance

For the random variables $X_1, X_2, \ldots, X_n$, we define the **sample mean** to be

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

and the **sample variance** to be

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

Properties of Sample Mean and Variance

For $X_1, X_2, \ldots, X_n$ IID, we have the following:

- $$\mathbb{E}[\bar{X}] = \mathbb{E}[X] = \mu$$
- $$Var(\bar{X}) = \frac{VarX}{n} = \frac{\sigma^2}{n}$$
- $$\mathbb{E}[S^2] = Var(X) = \sigma^2$$

Because the expected values of $\bar{X}$ and S^2 equal their target parameters, we say that they are **unbiased estimators** .

Binomial MLE Revisited

- Recall that with $X \sim Binomial(n, p)$, the MLE of p is

$$\hat{p} = \frac{X}{n}$$

- We have:

$$\mathbb{E}[\hat{p}] = \frac{X}{n} = \frac{np}{n} = p$$

$$Var(\hat{p}) = \frac{Var(X)}{n^2} = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

Moment Generating Functions

- The **moment generating function (mgf)** of a random variable X is

$$M(t) = \mathbb{E}[e^{tX}]$$

- Let $M^{(k)}(t)$ be the kth derivative of $M(t)$ with respect to t. We have that

$$M^{(k)}(0) = \mathbb{E}[X^k]$$

- The mgf can be used to compute **moments**, the expectations of the powers of the random variable:

$$E[X^k], \quad k = 1, 2, \ldots$$

- If $X_1, X_2, \ldots, X_n$ are independent and $Y = \sum_i X_i$, then

$$M_Y(t) = \prod_i M_{X_i}(t)$$

  where $M_{X_i}(t)$ is the mgf of X_i.

Example

- Let $X \sim Binomial(n, p)$. Again write $X = \sum_{i=1}^{n} X_i$, where the X_i are independent $Bernoulli(p)$ random variables.
- We have:

$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = pe^t + (1-p)(1) = pe^t + 1 - p$$

So

$$M_X(t) = \prod_i M_{X_i}(t) = (pe^t + 1 - p)^n$$

Example (Cont.)

Compute the mean and variance for a binomial random variable using the mgf:

$$M_X'(0) = np = \mathbb{E}[X]$$

and

$$M_X''(0) = np((n-1)p+1) = \mathbb{E}[X^2]$$

so

$$Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = np(1-p)$$