Exp and Variance

Thursday, January 30, 2020 1:23 PM

STA255_w...

STA255 Week-4

Shahriar Shams

27/01/2020

Review of Week-3

- Continuous random variable
- Probability density function (pdf)
  - A function from $\mathbb{R}$ to $\mathbb{R}$
  - For any continuous random variable X, $P(X = a) = 0$
  - Gives probability for an interval, $P(a \leq X \leq b) = \int_a^b f(x)dx$
- Distribution function (F)
  - for a continuous random variable, F is continuous every where
- Some common continuous distribution (Uniform, Exponential, Gamma, Chi-sq, Normal)
- Percentile of a distribution
  - $100 * p$th percentile, $\eta(p) \implies F(\eta(p)) = \int_{-\infty}^{\eta(p)} f(x)dx = p$
  - median is the 50th percentile of any distribution.

Learning goals

- Expectation
- Variance and standard deviation
- Moment generating function and it's use

Expectation

- A random variable and it's corresponding pdf/pmf contains all the information about the distribution or the underlying trial and outcomes of interest.
- Often instead of knowing the whole distribution we are interested in calculating summary measures
- *Expectation* is one of such summaries which is also called *center of the distribution* or *mean*
- Intuitively *Expectation* is the average value of the random variable.

For discrete random variable

Let's start with calculating simple average

- Task: what's the average of {1,1,1,1,2,2,2,2,3,3}
  - $\frac{1+1+1+1+2+2+2+2+3+3}{10} = \frac{(1*4)+(2*4)+(3*2)}{10} = 1 * \frac{4}{10} + 2 * \frac{4}{10} + 3 * \frac{2}{10}$
  - 4/10, 4/10 and 2/10 are the probabilities of selecting 1,2 and 3 (respectively) from the given list of numbers.
  - average is sum of the multiplication of numbers and their corresponding probabilities.

Definition of ***Expectation*** for discrete random variable

Suppose we have a discrete random variable X that takes value $x_1, x_2, x_3, ...$ with probability mass function, $p(x) = P(X = x)$. Then the *expectation* of X is defined as

$$E(X) = \sum_i x_i * P(X = x_i) = \sum_i x_i * p(x_i)$$

Example-1: Let X represents the number shown by a fair die. The corresponding pmf

x	1	2	3	4	5	6
$p(x)$	1/6	1/6	1/6	1/6	1/6	1/6

Then,

$$E(X) = 1 * 1/6 + 2 * 1/6 + ... + 6 * 1/6 = 3.5$$

Example-2: $X \sim Bern(p)$. $E(X) =?$

- X takes value 1 with probability p and value 0 with probability $1 - p$
- Hence, $E(X) = 1 * p + 0 * (1 - p) = p$

Example-3: $X \sim Pois(\lambda)$. $E(X) =$? (Exercise 104, page 152)

$$\begin{aligned}
E(X) &= \sum_{x=0}^{\infty} x * \frac{e^{-\lambda}\lambda^x}{x!} \\
&= 0.\frac{e^{-\lambda}\lambda^0}{0!} + \sum_{x=1}^{\infty} x * \frac{e^{-\lambda}\lambda^x}{x!} \\
&= 0 + \sum_{x=1}^{\infty} x * \frac{e^{-\lambda}\lambda^x}{x(x-1)!} \\
&= \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^x}{(x-1)!} \\
&= \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{(x-1)}\lambda}{(x-1)!} \\
&= \lambda \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{(x-1)}}{(x-1)!} \\
&= \lambda \sum_{y=0}^{\infty} \frac{e^{-\lambda}\lambda^{y}}{y!} \quad ; where, y = x-1 \\
&= \lambda * 1 = \lambda
\end{aligned}$$

Homework: Find $E[X(X-1)] = \sum_{x=0}^{\infty} x(x-1) * \frac{e^{-\lambda}\lambda^x}{x!}$ in the same way we found E[X]

$= \lambda^2$

For continuous random variable

Suppose we have a continuous random variable X with pdf f(x). The *expectation* of X is defined as

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

Example-1: $X \sim Exp(\lambda)$. $E(X) =$?

By definition,

$$E(X) = \int_{0}^{\infty} x\lambda e^{-\lambda x} dx$$

- Option-1: Use the "integration by parts" method that you have learned in calculus course and actually do the integration.
- Option-2: Use the *pdf* of a $Gamma(\alpha, \lambda)$ distribution to save some calculation.

pdf of $G(\alpha, \lambda)$

$= \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$

– calculate $\int_0^\infty x\lambda e^{-\lambda x}dx$

– calculate $\int_0^\infty x^2\lambda e^{-\lambda x}dx$

$$
\begin{aligned}
E[X] &= \int_0^\infty x \lambda e^{-\lambda x} dx \\
&= \lambda \int_0^\infty x e^{-\lambda x} dx \\
&= \lambda \int_0^\infty x^{2-1} e^{-\lambda x} dx \quad \Rightarrow \alpha = 2 \\
&= \frac{\Gamma(2)}{\lambda^2} \lambda \int_0^\infty \underbrace{\frac{\lambda^2}{\Gamma(2)} x^{2-1} e^{-\lambda x}}_{\text{pdf of } G(2,\lambda)} dx \\
&= \frac{\Gamma(2)}{\lambda^2} \lambda \\
&= \frac{1}{\lambda}
\end{aligned}
$$

$\Gamma(2) = 1$

Similarly,

$$
\begin{aligned}
E[X^2] &= \int_0^\infty x^2 \lambda e^{-\lambda x} dx \\
&= \cdots \\
&= \frac{2}{\lambda^2}
\end{aligned}
$$

Few properties of *Expectation*

- Expectation of any symmetric distribution is the point of symmetry
  - For example, if $X \sim N(\mu, \sigma^2)$ then $E(X) = \mu$
- Expectation of a constant (c) is that constant.
  - The idea is: average of 5,5,5,5,5,... is 5
  - $E(X)$ is a constant $\implies E(E(X)) = E(X)$
- Let g(X) be any real valued function of X
  - For discrete, $E(g(X)) = \sum_i g(a_i) * p(a_i)$
  - For continuous, $E(g(X)) = \int_{-\infty}^{\infty} g(x) f(x) dx$
- Expectation of a linear function of X is the function applied on the expectation of X
  - $E(aX + b) = aE(X) + b$, where a and b are constants.
  - Prove this...
  - Use of this property: if $X \sim N(\mu, \sigma^2)$ then $E(X - \mu) = 0$
  - In general, $E(X - E(X)) = 0$ for any X
- Expectation is a linear operator. Suppose we have random variables X and Y and constants r and s
  - $E(rX + sY) = rE(X) + sE(Y)$

Example: Suppose $X_1, X_2, ..., X_n$ are n independent $Bernoulli(p)$ variables. Let, $Y = X_1 + X_2 + ... + X_n$. We know by now that $Y \sim Bin(n, p)$. $E(Y) = ?$

$$E(Y) = E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n) = p + p + ... + p = np$$

Recall: In week-2 we learned, $Bin(n, p) \to Pois(\lambda)$ where parameter of the Poisson distribution λ satisfies $\lambda = np$ which is just the mean of the distribution.

Note: Along with *Expectation*, median or mode (the value with the highest pmf or pdf) are also used as measure of the center of the distribution.

end of day-1

$X \sim Bern(p)$

$E[X] = 1 * p + 0 * (1-p)$

$E[X^2] = 1^2 * p + 0^2 * (1-p)$

$E[X^3] = 1^3 * p + 0^3 * (1-p)$

$X \sim Unif[\alpha, \beta]$

$f(x) = \frac{1}{\beta - \alpha}$

$E[X^2] = \int x^2 * \frac{1}{\beta-\alpha} dx$

$E[aX+b]$

$= \int (ax+b) f(x) dx$

$= \int ax f(x) dx + \int b f(x) dx$

$= a \underbrace{\int x f(x) dx}_{E[X]} + b \underbrace{\int f(x) dx}_{1}$

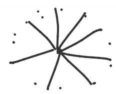
$= a E[X] + b$

$X \sim N(\mu, \sigma^2)$

$E\left[\frac{X-\mu}{\sigma}\right] = E\left[\frac{1}{\sigma} X - \frac{\mu}{\sigma}\right] = \frac{1}{\sigma} E[X] - \frac{\mu}{\sigma}$

$= 0$

Variance



- *Variance* is another commonly used and important summary measure.
- Remember, *Expectation* / *mean* is the center of the distribution.
- *Variance* is the **spread**(around the mean) of the distribution.
  - In naive words, it tells us how much variation we have in our data (random variable)
- This is also an expectation but not of X rather it's a expectation of a function of X
- $(X - E(X))^2$ is the squared distance between any X from it's own mean $E(X)$
- The expectation of these squared distances are called variance.

Definition of *variance*

The *variance* of a random variable X denoted by $Var(X)$ or simply $V(X)$ is defined as

$$V(X) = E[(X - E(X))^2]$$

Example-1: For the faces of a die (table given in the bottom of page 2 of this pdf) we got $E(X) = 3.5$. Variance,

$$V(X) = (1 - 3.5)^2 * 1/6 + (2 - 3.5)^2 * 1/6 + ... + (6 - 3.5)^2 * 1/6 = 2.916667$$

Example-2: $X \sim Bern(p)$, we already know $E(X) = p$, then $V(X) =?$

$$V(X) = E[(X - E(X))^2] = E[(X - p)^2] = (1 - p)^2 * p + (0 - p)^2 * (1 - p) = p(1 - p)$$

Re-writing the formula of *variance*

$$\begin{aligned}
\operatorname{Var}(X) &= \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] \\
&= \int_{-\infty}^{\infty} (x - \mathrm{E}[X])^2 f(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \left(x^2 - 2x\mathrm{E}[X] + (\mathrm{E}[X])^2\right) f(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x - 2\mathrm{E}[X]\int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x + (\mathrm{E}[X])^2 \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x \\
&= \mathrm{E}\left[X^2\right] - 2(\mathrm{E}[X])^2 + (\mathrm{E}[X])^2 \\
&= \mathrm{E}\left[X^2\right] - (\mathrm{E}[X])^2.
\end{aligned}$$

E[X] is a const.

Example-3: If $X \sim Pois(\lambda)$, Show that $V(X) = \lambda$

$$\begin{aligned} V(X) &= E[X^2] - (E[X])^2 \\ &= E[X(X-1)] + E[X] - (E[X])^2 \\ &= \lambda^2 + \lambda - \lambda^2 \\ &= \lambda \end{aligned}$$

Note: For $Pois(\lambda)$, $\boxed{E[X] = V[X] = \lambda}$

Example-4: $X \sim Exp(\lambda)$, $V(X) = ?$

$E[X] = \frac{1}{\lambda}$, $E[X^2] = \frac{2}{\lambda^2}$

$\Rightarrow V[X] = E[X^2] - (E[X])^2$

$= \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$

Properties of *variance*

- Variance of a constant is zero.
  - There is not variation in {5,5,5,. . . ,5}. So variance = 0

- For a random variable X and constants r and s

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

$V[aX+b] = E\left[\left\{(aX+b) - E(aX+b)\right\}^2\right]$

$=$

$= a^2 V[X]$

  - $X \sim N(\mu, \sigma^2)$. Show $V(Z) = 1$ where $Z = \frac{X-\mu}{\sigma}$

- Let X and Y are two **independent** random variables. Then

$$V(X + Y) = V(X) + V(Y)$$

Example: Suppose $X_1, X_2, ..., X_n$ are n **independent** $Bernoulli(p)$ variables. Let, $Y = X_1 + X_2 + ... + X_n$. We know by now that $Y \sim Bin(n, p)$. $V(Y) = ?$

if $X \sim Bern(p)$

$V[X] = p(1-p)$

$$V(Y) = V(X_1+X_2+...+X_n) = V(X_1)+V(X_2)+...+V(X_n) = p(1-p)+p(1-p)+...+p(1-p) = np(1-p)$$

Standard Deviation

Standard deviation denoted by sd defined as the positive square root of the variance.

$$sd(X) = \sqrt{V(X)}$$

Expectation, variance or sd of a list of numbers using R

Example using the faces of a die.

```r
x=c(1,2,3,4,5,6)

#Expectation of X
m=mean(x)
m
```

```
## [1] 3.5
```

```r
#variance of
v= mean( (x-m)^2 )
v
```

```
## [1] 2.916667
```

```r
#standard deviation of x
sqrt(v)
```

```
## [1] 1.707825
```

Moments and Momets Generating function

Moments

- Expected values of integer power of the random variable X or $X - \mu$ are called moments where $\mu = E[X]$.
- $E[X^r]$ is called rth moment around 0.
  - In statistical texts it's often called *raw moments*.
  - $E[X]$ or mean is the first moment around zero/ first raw moment.
- $E[(X - \mu)^r]$ is called the rth moment around μ.
  - In statistical texts it's often called *central moments*.
  - $V[X]$ or variance is the second moment around μ/ second central moments.
- Third central moment, $E[(X - \mu)^3]$ measures the skewness of the distribution.

Moment generating funciton (MGF)

- Calculating all these individual moments are often troublesome.
- Instead of calculating individual moments often we calculate what is called the moment generating function, $M_X(t)$

$$M_X(t) = E[e^{tx}]$$

$E[e^{tX}]$

$$E[e^{tX}] = \int_x e^{tx} f(x)dx$$

$$\frac{d}{dt}E[e^{tX}] = \frac{d}{dt}\int_x e^{tx} f(x)dx$$

$$= \int_x \frac{d}{dt}e^{tx} f(x)dx$$

$$= \int_x x e^{tx} f(x)dx$$

$$t = 0 \Rightarrow = \int_x x f(x)dx = E[X]$$

Differentiating twice,

$$\frac{d^2}{dt^2}E[e^{tX}] = \int_x \frac{d^2}{dt^2} e^{tx} f(x)dx$$

$$= \int_x x^2 e^{tx} f(x)dx$$

$$t = 0 \Rightarrow \int_x x^2 f(x)dx = E[X^2]$$

- Differentiating this function r times and then putting $t=0$ gives us the rth moment around 0/raw moment.

$$E[X^r] = M^{(r)}(0) = \frac{d^r}{dt^r} M_X(t)\Big|_{t=0}$$

Example: For Bernoulli(p), $M_X(t) = (1-p) + pe^t$ (page 122 of the text book)

Homework:

- differentiate this function once with respect to t and then put $t=0$. That should give you $E[X]$
- differentiate again, and then put $t=0$. That should give you $E[X^2]$

Few important properties of MGF

- Suppose we have a random variable $Y = aX + b$ which is a linear function of X. Then

$$M_Y(t) = e^{bt} M_X(at)$$

$$M_Y(t) = E[e^{tY}] = E[e^{t(aX+b)}]$$
$$= E[e^{taX} . e^{tb}]$$
$$= e^{tb} E[e^{taX}]$$
$$= e^{bt} M_X(at)$$

(show)

- Suppose we have two random variables X and Y that are independent. Let $Z = X + Y$. Then

$$M_Z(t) = M_X(t) * M_Y(t)$$

MGF uniqely characterizes a distribution

- If two random variables produces the same moment generating function that means they follow the same distribution.
- The constants in the MGF are parameters of the distribution.
- Lets use this with few of the previous properties to identify distributions.

Example-1: Suppose $X \sim Pois(\lambda_1)$, $Y \sim Pois(\lambda_2)$ and X and Y are independent. Further you are told that the MGF of a $pois(\lambda)$ is $e^{\lambda}(e^t - 1)$ (proof on page 149).

wrong $e^{\lambda(e^t-1)}$

Show that $X + Y \sim Pois(\lambda_1 + \lambda_2)$

Example-2: Suppose $X \sim N(\mu, \sigma^2)$ and $M_X(t) = e^{\mu t + \sigma^2 t^2/2}$. Find the MGF of $Z = (X - \mu)/\sigma$

Homework

From the exercise

Chapter 3.3

28, 31, 35, 36, 37

Chapter 3.4

49(a), 51, 52, 56(a)

Homework

From the exercise

Chapter 3.3

28, 31, 35, 36, 37

Chapter 3.4

49(a), 51, 52, 56(a)

Chapter 3.5

70, 72

Chapter 4.2

20, 25, 26, 29, 30, 37

Example: 2

$$Z = \frac{1}{\sigma} X - \frac{\mu}{\sigma} \Rightarrow a = \frac{1}{\sigma}, b = -\frac{\mu}{\sigma}$$

$$Z = aX + b$$

$$M_Z(t) = e^{bt} M_X(at)$$

$$= e^{-\frac{\mu}{\sigma} t} * e^{\mu * \frac{1}{\sigma} t + \sigma^2 \frac{1}{\sigma^2} t^2/2}$$

$$= e^{t^2/2} = e^{0*t + 1 \cdot t^2/2}$$

$$Z \sim N(0,1)$$

Example: 1

$$X \sim Poin(\lambda_1) \quad Y \sim Poin(\lambda_2)$$

$$M_X(t) = e^{\lambda_1 (e^t - 1)} \quad e^{\lambda_2 (e^t - 1)}$$

$$Z = X + Y$$

$$M_Z(t) = M_X(t) * M_Y(t)$$

$$= e^{\lambda_1 (e^t - 1)} * e^{\lambda_2 (e^t - 1)}$$

$$= e^{\lambda_1 (e^t - 1) + \lambda_2 (e^t - 1)}$$

$$= e^{(\lambda_1 + \lambda_2)(e^t - 1)}$$

$$Z \sim Pois(\lambda_1 + \lambda_2)$$