Stat 343, Fri 18-Sep-2020 -- Fri 18-Sep-2020

Probability and Statistics

Fall 2020

Friday, September 18th 2020

Wk 3, Fr

Topic:: mean, variance of random variable

σ

mean μ measures center of a distribution

Definition 1 (2.5.7): Let X be a discrete r.v. The **variance** of X, denoted by $\mathrm{Var}(X)$ or by σ_X^2, is that variables mean squared deviation from the mean. More explicitly, that is

$$\mathrm{Var}(X) \;=\; \mathrm{E}((X-\mu_X)^2).$$

pmf

x	0	1	2
$f_X(x)$	$(0.7)^2$	$\binom{2}{1}(0.3)(0.7)$	$(0.3)^2$

Example. Compute by hand the variance for X when

(a) $X \sim \mathrm{Binom}(2, 0.3)$

$$\mu_X = n\pi = 0.6$$

$$\sigma_X^2 = (0-0.6)^2(0.7)^2 + (1-0.6)^2 \cdot (2\cdot(0.3)(0.7)) + (2-0.6)^2(0.3)^2$$

(b) $X \sim \mathrm{Binom}(3, 0.5)$

Theorem 1 (2.5.8): Let X be a discrete r.v. Then $\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$.

$$\mathrm{Var}(X) = E((X-\mu_X)^2) = E(X^2 - 2\mu_X X + \mu_X^2)$$

$$= E(X^2) - E(2\mu_X X) + E(\mu_X^2)$$

$$= E(X^2) - 2\mu_X \underbrace{E(X)}_{\mu_X} + \mu_X^2$$

$$= E(X^2) - 2\mu_X^2 + \mu_X^2 = E(X^2) - \mu_X^2$$

Example: From Problem B.21 we have the pmf

x	0	1	2	3	4
$f(x)$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{12}$

Last time, we found $E(X) = 5/3$ and $E(X^2) = 25/6$. Use these to calculate $\text{Var}(X)$.

$$\text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{25}{6} - \left(\frac{5}{3}\right)^2$$

Example: A brief survey

marginal totals

# X parents who attended Calvin	Planned on Calvin full year prior ($Y = 0$)	Later choice ($Y = 1$)	
0	4	6	10
1	1	1	2
2	2	5	7
	7	12	19

Speak of $f_{X,Y}(x,y)$ gives joint distribution of X, Y.

$$= P(X=1 \text{ and } Y=0) = 1/19 = f_{X,Y}(1,0) \text{ - joint pmf}$$

$$P(X=2 \text{ and } Y=1) = 5/19 = f_{X,Y}(2,1)$$

$$P(X=2) = \frac{7}{19} = f_X(2) \quad \text{marginal distribution of } X.$$

$$P(Y=0) = 7/19 = f_Y(0) = \sum_x f_{X,Y}(x,0)$$

$$P(X=0 \mid Y=1) = 6/12 = f_{X \mid Y=1}(0) = \frac{f_{X,Y}(0,1)}{f_Y(1)}$$

$$= \frac{6/19}{12/19} = \frac{6}{12}$$

$P(X=0, Y=1) = 6/19$

$\stackrel{?}{=} P(X=0) \cdot P(Y=1)$

$= \left(\frac{10}{19}\right)\left(\frac{12}{19}\right)$ No.

starting value fixed from class (I used $P(X=0 \mid Y=1)$ which was incorrect!)

Joint Distributions

Our calculation of probabilities has led to the consideration of the concurrence of two events—get a "spade" and a "king", roll "doubles" and a "number larger than 6", etc. And, as many events are depicted with random variables, this naturally leads to considering two or more random variables together. To facilitate answering questions such as $P(2 \leq X \leq 4 \text{ and } Y = 5)$, we would like to have (in the case where X, Y are discrete r.v.s) a **joint pmf**, a function that yields values

$$f_{X,Y}(x,y) = P(X = x \,\text{and}\, Y = y) \qquad \text{abbreviated as } P(X = x, Y = y).$$

Naturally, the idea can be extended to that for a joint pmf of k discrete r.v.s. If one has such a joint pmf, one easily recovers the individual (or **marginal**) distributions for X, Y:

$$f_X(x) = \sum_y f_{X,Y}(x,y) \qquad \text{and} \qquad f_Y(y) = \sum_x f_{X,Y}(x,y).$$

We can also obtain conditional distributions

$$f_{X|Y=y}(x) \;=\; P(X = x|Y = y) \;=\; \frac{P(X = x, Y = y)}{P(Y = y)} \;=\; \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

(similar definition for $f_{Y|X=x}(y)$).

Definition 2: Suppose f is the joint pmf of discrete r.v.s X, Y, and let $t\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. Then $Z = t(X,Y)$ is a discrete r.v. with pmf given by

$$P(Z = z) \;=\; \sum_{\left\{(x,y)\,\middle|\,t(x,y)=z\right\}} f(x,y) \;=:\; \sum_{t(x,y)=z} f(x,y).$$

Definition 3: Suppose f is the joint pmf of discrete r.v.s X, Y. We say X, Y are **independent** if for every x and y,

$$f(x,y) \;=\; f_X(x) \cdot f_Y(y).$$

Have a great weekend!

Theorem 2: Let X, Y be discrete random variables. Then

(i) $\mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y)$.

(ii) $\mathrm{E}(XY) = \mathrm{E}(X) \cdot \mathrm{E}(Y)$, if X and Y are independent.

(iii) $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$, if X and Y are independent.

Proof:

part (iii):
We have

$$\begin{aligned}
\mathrm{Var}(X+Y) &= \mathrm{E}((X+Y)^2) - [\mathrm{E}(X+Y)]^2 \;=\; \mathrm{E}(X^2 + 2XY + Y^2) - [\mathrm{E}(X+Y)]^2 \\
&= \mathrm{E}(X^2) + \mathrm{E}(2XY) + \mathrm{E}(Y^2) - [\mathrm{E}(X)]^2 - 2\,\mathrm{E}(X)\,\mathrm{E}(Y) - [\mathrm{E}(Y)]^2 \\
&= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 + \mathrm{E}(Y^2) - [\mathrm{E}(Y)]^2 + 2\,\mathrm{E}(XY) - 2\,\mathrm{E}(X)\,\mathrm{E}(Y) \\
&= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2[\mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)] \\
&= \mathrm{Var}(X) + \mathrm{Var}(Y),
\end{aligned}$$

by part (ii) of the theorem. □

Expected values and variances revisited

The results in the last theorem give us the tools for computing means and variances for several standard statistical models.

Binomial distributions

Special case: $X \sim \mathsf{Binom}(1, \pi)$. Such an X called a **Bernoulli random variable**. Here

$$\mu =$$

$$\mathrm{Var}(X) =$$

General case: $X \sim \mathsf{Binom}(n, \pi)$. Note

$$X = X_1 + X_2 + \cdots + X_n,$$

with each $X_j \sim \mathsf{Binom}(1, \pi)$ (Bernoulli) and the collection $X_1, \ldots, X_n$ is independent in the sense that

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = P(X_1 = x_1) \cdot P(X_2 = x_2) \cdots P(X_n = x_n).$$

By an analog to the last theorem,

$$\mathrm{E}(X) =$$

$$\mathrm{Var}(X) =$$

Negative binomial distributions. In similar fashion, an $X \sim \mathsf{NBinom}(n, \pi)$ may be thought as the sum of independent random variables $X = X_1 + X_2 + \cdots + X_n$ where each $X_j \sim \mathsf{NBinom}(1, \pi)$. ($X_j$ counts the number of failed attempts between the $(j-1)^{\text{st}}$ success and the j^{th} one.) We have not previously calculated the variance of a geometric r.v., but Pruim calculated the mean, on p. 79, to be $(1-\pi)/\pi$. Thus,

$$\mathrm{E}(X) = \sum_{j=1}^{n} \mathrm{E}(X_j) =$$

The sum of independent, identically distributed r.v.s. In both the binomial and negative binomial cases above, we could write $X = X_1 + X_2 + \cdots + X_n$, where all the X_js come from the same distribution, and the collection $X_1, \ldots, X_n$ is independent. We abbreviate these assumptions about $X_1, \ldots, X_n$ by calling the i.i.d. random variables, where i.i.d. stands for *independent and identically distributed*. If μ, σ^2 stand for the mean and variance, respectively, of the distribution common to the X_j, then their sum has mean and variance

$$\mathrm{E}\,(X) \;=\; \sum_{j=1}^{n} \mu \;=\; n\mu,$$

and

$$\mathrm{Var}(X) \;=\; \sum_{j=1}^{n} \sigma^2 \;=\; n\sigma^2.$$