2) Random Variables

Vitor Kamada

June 2018

Tables, Graphics, and Figures from

Introductory Statistics with

Randomization and Simulation

Diez et al. (2014): APPENDIX A3 - Probability

Random Variable: Modelling

Two books are assigned for a statistics class: a textbook and its corresponding study guide. The bookstore determined 20% of enrolled students do not buy either book, 55% buy the textbook only, and 25% buy both books.
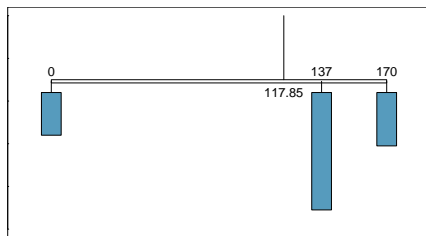
The textbook costs $137 and the study guide $33.

How much revenue should the bookstore expect from this class of 100 students?

Expectation: $\mu = E(X) = \sum\limits_{i=1}^{k} x_i P(X = x_i)$

i	1	2	3	Total
x_i	$0	$137	$170	–
$P(X = x_i)$	0.20	0.55	0.25	1.00

$$x_1 P(X = x_1) + ... + x_k P(X = x_k)$$

$$= 0(.20) + 137(.55) + 170(.25) = \$117.85$$

Variance and Standard Deviation

$$\sigma^2 = \sum_{j=1}^{k} (x_j - \mu)^2 P(X = x_j)$$

$$= (x_1 - \mu)^2 P(X = x_1) + ... + (x_k - \mu)^2 P(X = x_k)$$

i	1	2	3	Total
x_i	\$0	\$137	\$170	
$P(X = x_i)$	0.20	0.55	0.25	
$x_i \times P(X = x_i)$	0	75.35	42.50	117.85
$x_i - \mu$	-117.85	19.15	52.15	
$(x_i - \mu)^2$	13888.62	366.72	2719.62	
$(x_i - \mu)^2 \times P(X = x_i)$	2777.7	201.7	679.9	3659.3

$$\sigma = \sqrt{3659.3} = \$60.49$$

Linear Combinations of Random Variables

$$Z = aX + bY$$

$$E(Z) = aE(X) + bE(Y)$$

Leonard has invested \$6000 in Google and \$2000 in Exxon Mobil. Suppose Google and Exxon Mobil stocks have recently been rising 2.1% and 0.4% per month, respectively.

$$E(6000X + 2000Y)$$

$$6000(.021) + 2000(.004) = \$134$$

Variability of Independent Random Variables

	Mean ($\bar{x}$)	Standard deviation (s)	Variance (s^2)
GOOG	0.0210	0.0846	0.0072
XOM	0.0038	0.0519	0.0027

$$Var(aX + bY) = a^2 Var(X) + b^2 Var(Y)$$

$$Var(6000X + 2000Y)$$

$$= 6000^2(.0072) + 2000^2(.0027) = 270,000$$

$$SD(6000X + 2000Y) = \sqrt{270,000} = \$520$$

Covariance of Random Variables

$$\sigma_{xy} = Cov(X, Y) = E[(X - \mu_x)(Y - \mu_y)]$$

$$= E(XY - X\mu_y - \mu_x Y + \mu_x \mu_y)$$

$$= E(XY) - \mu_y E(X) - \mu_x E(Y) + \mu_x \mu_y$$

$$= E(XY) - \mu_x \mu_y$$

If X and Y are independent, then $Cov(X, Y) = 0$

Variance for Dependent Random Variables

$$Var(X) = E[(X - \mu_x)^2]$$

$$Var(X + Y) = E[(X + Y - \mu_x - \mu_y)^2]$$

$$= E[(X - \mu_x + Y - \mu_y)^2]$$

$$= E[(X - \mu_x)^2 + (Y - \mu_y)^2 + 2(X - \mu_x)(Y - \mu_y)]$$

$$= Var(X) + Var(Y) + 2Cov(X, Y)$$

Correlation for Random Variable

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X) \times Var(Y)}} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

$$-1 \leq \rho \leq 1$$

Suppose X can take only the three values -1, 0, and 1, and that each of these three values has the same probability. Let $Y = X^2$

$$E(XY) = E(X^3) = E(X) = 0$$

$\therefore X$ and Y are dependent, but uncorrelated

$$E(X) = x_1 p_1 + x_2 p_2 + ... + x_k p_k$$

$$E(cX) = cx_1 p_1 + cx_2 p_2 + ... + cx_k p_k$$

$$= c(x_1 p_1 + x_2 p_2 + ... + x_k p_k)$$

$$= cE(X)$$

$$Var(X) = E[(X - \mu_x)^2]$$

$$Var(cX) = E[(cX - c\mu_x)^2]$$

$$= E[c^2(X - \mu_x)^2]$$

$$= c^2 E(X - \mu_x)^2$$

$$= c^2 Var(X)$$