5 Multivariate Probability distributions

5.1 *Introduction*

- A specific set of outcomes, or sample measurements, may be expressed in terms of the intersection of the n events ($Y_1=y_1$), ($Y_2=y_2$),… ,($Y_n=y_n$), which we will denote as ($Y_1=y_1$, $Y_2=y_2$,… ,$Y_n=y_n$), or as ($y_1,y_2,\dots,y_n$)

5.2 *Bivariate and multivariate probability distributions*

- Many random variables can be defined over the same sample space.
- E.g. tossing a pair of dice.
  - The sample space contains ...equiprobable sample points
  - Could consider random variables such as
    Y1: # of dots appearing on die 1
    Y2: # of dots appearing on die 2
    Y3: the sum of the # of dots on the dice
    Y4: the product of the # of dots on the dice
  - 36 numerical events (y_1,y_2) are all equally likely and thus are assigned the sample probability of ???
  - The bivariate probability function is
    $p(y_1,y_2)=P(Y_1=y_1,Y_2=y_2)=1/36$, $y_1=1,\dots,6$, $y_2=1,\dots,6$
  - A graph? A theoretical, three-dimensional relative frequency histogram for the pairs of observations (y_1,y_2)
- Definition 5.1 Let Y1 and Y2 be discrete RVs. The "joint (or bivariate) probability distribution" for Y1 and Y2 is given by
  $p(y_1,y_2)=P(Y_1=y_1,Y_2=y_2)$, $-\infty<y_1<\infty$, $-\infty<y_2<\infty$
  The function $p(y_1,y_2)$ is called the "joint probability function"
- *Theorem 5.1 If Y1 and Y2 are discrete RVs with joint probability function p(y1,y2), then*
  1. $p(y_1,y_2)\geq 0$ for all y_1, y_2
  2. $\sum_{y_1,y_2} p(y_1,y_2)=1$, where the sum is over all values (y1,y2) that are assigned nonzero probabilities
- Example 5.1 Three checkout counters. Two customers choose a counter at random independently. Let Y1 = # of customers who choose counter 1 and Y2 = # who choose counter 2. The joint distribution of Y1 and Y2? (table)
- Exercises 5.1-5.16

- Keywords: joint probability distribution/function

5.3 Marginal and conditional probability distributions

- Reconsider Y1 and Y2 in 2-dice tossing experiment. What're $P(Y_1=y_1)$ where y1=1,…,6?
- $P(Y_1 = y_1) = p_1(y_1) = \sum_{y_2=1}^{6} p(y_1, y_2)$ and similarly for $p_2(y_2)$
- Definition 5.4 Let Y1 and Y2 be jointly discrete RVs with probability function $p(y_1,y_2)$. Then the "marginal probability functions" of Y1 and Y2, respectively, are given by
  $p_1(y_1) = \sum_{y_2} p(y_1, y_2)$ and $p_2(y_2) = \sum_{y_1} p(y_1, y_2)$
    - Why is it called "marginal"?
- Example 5.5 Choosing a committee of two people from 3 Republicans, 2 Democrats and 1 independent. Y1=# of republicans on the committee and Y2=# of Democrats on the committee. The joint probability dist of Y1 and Y2? The marginal distribution of Y1?
- Recall the multiplicative law: $P(A\cap B)=P(A)P(B|A)= P(B)P(A|B)$.
  Letting $A=(Y_1=y_1)$ and $B=(Y_2=y_2)$, we get
  …
- Definition 5.5 if Y1 an Y2 are jointly discrete RVs with joint probability function p(y1,y2) and marginal probability functions p1(y1) and p2(y2), respectively, then the "conditional discrete probability function" of Y1 given Y2 is
  $$p(y_1 | y_2) = P(Y_1 = y_1 | Y_2 = y_2) = \frac{P(Y_1 = y_1, Y_2 = y_2)}{P(Y_2 = y_2)} = \frac{p(y_1, y_2)}{p_2(y_2)}$$
  provided that p2(y2)>0 (it is undefined if p2(y2)=0)
- Example 5.7 (5.5 continued) Conditional distribution $P(Y_1=y_1|Y_2=1)$=?
- Exercises 5.17-5.36
- Keywords: marginal and conditional probability distribution/function

5.4 Independent Random Variables

- Two events A and Be are independent if $P(A\cap B)=P(A)P(B)$. If Y1 and Y2 are independent, we would like to have P(Y1∈[a,b], Y2∈[c,d]) = P(Y1∈[a,b]) P(Y2∈[c,d]) for any choice of real numbers a,b,c, and d.
- Theorem 5.4 (Definition) Two discrete RVs Y1 an Y2 are independent iff
  p(y1,y2)=p1(y1)p2(y2)
  for all pairs of real numbers (y1,y2).
- Example 5.9 For the die-tossing problem, show that Y1 and Y2 are independent

- Example 5.10 In Example 5.5 (Republicans and Democrats), is the # of republicans in the sample independent of the # of Democrats?
- Exercises 5.37-5.61
- Keywords: independence

5.5 The expected value of a function of random variables

- Definition 5.9 Let g(y1,…,Yk) be a function of the discrete RVs Y1,…,Yk which have probability function p(y1,…,yk). Then the "expected value" of g(Y1,…,Yk) is
  $E[g(Y1,\ldots,Yk)] = \sum_{yk}\ldots\sum_{y2}\sum_{y1} g(y1,\ldots,yk)p(y1,\ldots,yk)$

5.6 Special theorems

- *Theorem 5.6 For any constant c, E(c)=c.*
- *Theorem 5.7. For $g(Y_1,Y_2)$ a function of two RVs and a constant c, $E[cg(Y_1,Y_2)] = cE[g(Y_1,Y_2)]$.*
- *Theorem 5.8. For functions $g_1(Y_1,Y_2), g_2(Y_1,Y_2),\ldots,g_k(Y_1,Y_2)$ of Y_1,Y_2, $E[g_1(Y_1,Y_2)+g_2(Y_1,Y_2)+\ldots+g_k(Y_1,Y_2)]=E[g_1(Y_1,Y_2)]+E[g_2(Y_1,Y_2)]+\ldots+E[g_k(Y_1,Y_2)]$.*
- *Theorem 5.9 If two RVs Y_1,Y_2 are independent and $g(Y_1)$ and $h(Y_2)$ are functions of only Y_1 and Y_2 respectively, then $E[g(Y_1)h(Y_2)]=E[g(Y_1)]\,E[g(Y_2)]$*
- Exercises 5.62-5.74
- Keywords: E[g(Y1,…,Yk)]

5.7 The covariance of two random variables

- Dependence and the linear relationship between two variables (some plots of either rho=0 and rho>0). Some motivation.
- Definition 5.10 If Y1 and Y2 are RVs with means μ1 and μ2, respectively, the "covariance" of Y1 and Y2 is given by
  $$Cov(Y_1,Y_2) = E[(Y_1-\mu_1)(Y_2-\mu_2)]$$
  The "correlation coefficient" ρ is a standardized covariance defined as
  $\rho = \frac{Cov(Y_1,Y_2)}{\sigma_1\sigma_2}$ and is easier to interpret. (why?)
- *Theorem 5.10* $Cov(Y_1,Y_2) = E(Y_1Y_2) - E(Y_1)E(Y_2)$.
  - Proof
- *Theorem 5.11 If Y1 and Y2 are independent RVs, then* $Cov(Y_1,Y_2) = 0$.
- Example 5.24 Dependent but zero correlation (implication?)
- Exercises 5.75-5.85
- Keywords: Covariance, correlation

5.8 The expected value and variance of linear functions of random variables

- *Theorem 5.12 Let Y1,…,Yn and X1,…,Xm be RVs with $E(Y_i)=\mu_i$ and $E(X_j)=\xi_j$. Define* $U_1 = \sum_{i=1}^{n} a_i Y_i$ *and* $U_2 = \sum_{j=1}^{m} b_j X_j$ *for constants $a_1,\ldots,a_m$ and $b_1,\ldots,b_m$. Then the following hold:*

  a $E(U_1) = \sum_{i=1}^{n} a_i \mu_i$

  b $V(U_1) = \sum_{i=1}^{n} a_i^2 V(Y_i) + 2\sum_{i<j} a_i a_j Cov(Y_i, Y_j)$

  c $Cov(U_1, U_2) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j Cov(Y_i, X_j)$
- Example 5.25 E(Yj)=1, 2, -1 and V(Yj)=1,3,5 and Cov(Y1,Y2)=-.4, Cov(Y1,Y3)=1/2 and Cov(Y2,Y3)=2. What's E(U), V(U), Cov(U,W) for $U=Y_1-2Y_2-Y_3$ and $W=3Y_1+Y_2$?
- Example 5.27 Let Y1,..,Yn~iid E(Yi)=μ and V(Yi)=σ^2. Define $\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$. What are $E(\overline{Y}), V(\overline{Y})$?
- Example 5.28 # of defective Y~Bin(10,p). What's the expected value and variance of the estimator of the fraction defective in the lot $\hat{p} = \frac{Y}{n}$.
- Exercises 5.86-5.98
- Keywords: Linear functions of RVs