Handout 9: Math/Stat 394: Probability I

Expectation Summary

Wellner; 2/23/2000

Definition 1. Random outcome X, Y yields payoff $g(X, Y)$ with average value

$$Eg(X,Y) \equiv \begin{cases} \sum\sum_{all x,y} g(x,y) p_{X,Y}(x,y) & \text{(discrete rv's)} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dx\,dy & \text{(continuous rv's)}\,. \end{cases}$$

Example 1.
$\mu_X \equiv EX$ and $\sigma_X^2 \equiv Var(X) \equiv E(X - EX)^2$.
$\mu_Y \equiv EY$ and $\sigma_Y^2 \equiv Var(Y) \equiv E(Y - EY)^2$.
μ_X and σ_X measure "center" and "spread" of X;
μ_Y and σ_Y measure "center" and "spread" of Y;

$\sigma_{X,Y} \equiv Cov(X, Y) \equiv E[(X - EX)(Y - EY)]$
and $\rho_{X,Y} \equiv Corr(X, Y) \equiv \sigma_{X,Y}/(\sigma_X \sigma_Y)$
measure the relationship between X and Y.

Theorem 1.

$E(a) = a$	
$E(aX+b) = aE(X) + b$	or $\mu_{aX+b} = a\mu_X + b$
$E(aX+bY+c) = aE(X) + bE(Y) + c$	or $\mu_{aX+bY+c} = a\mu_X + b\mu_Y + c$ always
$E(X+Y) = E(X) + E(Y)$	or $\mu_{X+Y} = \mu_X + \mu_Y$ always
$Var(aX+b) = a^2 Var(X)$	or $\sigma_{aX+b} = \|a\|\sigma_X$
$Var(X) = 0$ iff $P(X = \mu_X) = 1$	
$Var(X+Y) = Var(X) + Var(Y)$ if X and Y are independent	or $\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$ if X and Y are independent
$Var(X+Y)$ $= Var(X) + Var(Y) + 2Cov(X,Y)$ always	or $\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y + 2\sigma_{X,Y}$ always
$Var(aX+bY+c)$ $= a^2Var(X) + b^2Var(Y) + 2abCov(X,Y)$	or $\sigma^2_{aX+bY+c} = a^2\sigma^2_X + b^2\sigma^2_Y + 2ab\sigma_{X,Y}$

Computing Formulas:

$Var(X) = E(X^2) - (EX)^2$	or $\sigma^2_X = E(X^2) - \mu^2_X$
$Cov(X,Y) = E(XY) - EX \cdot EY$	or $\sigma_{X,Y} = E(XY) - \mu_X\mu_Y$
$Var(X) = Cov(X,X)$	
$Cov(X,Y) = 0$ whenever X and Y are independent	

The **key fact** in all this:
$E\{g(X)h(Y)\} = E\{g(X)\} \cdot E\{h(Y)\}$ for any $g(\cdot)$ and $h(\cdot)$ whenever X and Y are independent.
Thus $Cov(X,Y) = E\{(X-\mu_X)(Y-\mu_Y)\} = E(X-\mu_X)\cdot E(Y-\mu_Y) = 0\cdot 0 = 0$ when X and Y are independent.

$$Cov[\sum_{i=1}^{m} a_i X_i, \sum_{j=1}^{n} b_j Y_j] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j Cov[X_i, Y_j]$$

$$Var[\sum_{i=1}^{n} a_i X_i] = \sum_{i=1}^{n} a_i^2 Var[X_i] + \sum\sum_{i\neq j} a_i a_j Cov[X_i, X_j]$$

$$Var[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} Var[X_i] + \sum\sum_{i\neq j} Cov[X_i, X_j]$$

$$= nVar(X_1) + n(n-1)Cov[X_1, X_2]$$

if all variances and covariances are equal.

$$Var[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} Var(X_i) \qquad \text{if the } X_i's \text{ are independent}$$
$$= nVar(X_1) \qquad \text{if all the variances are equal}.$$

Example 2. (Urn mean and variance). An urn contains N balls numbered $a_1, \ldots, a_N$. Let X denote the $a-$value of one ball drawn at random from the urn. Then

$$\mu_X = \overline{a}_N \equiv \frac{1}{N}\sum_{i=1}^{N} a_i$$

and

$$\sigma_X^2 = \sigma_a^2 \equiv \frac{1}{N}\sum_{i=1}^{N}(a_i - \overline{a}_N)^2 = \frac{1}{N}\sum_{i=1}^{N} a_i^2 - \overline{a}_N^2 .$$

Example 3. (Sampling *with* replacement from an urn.) Let $T_n \equiv X_1 + \cdots + X_n$. Then

$$E(T_n) = n\overline{a}_N \qquad \text{and} \qquad Var(T_n) = n\sigma_a^2 ,$$

and hence for $\overline{X}_n \equiv T_n/n$,

$$E(\overline{X}_n) = \overline{a}_N \qquad \text{and} \qquad Var(\overline{X}_n) = \frac{1}{n}\sigma_a^2 .$$

Note that $Cov(X_1, X_2) = 0$ since the X_i's are independent.

Example 4. (Sampling *without* replacement from an urn.) Let $T_n \equiv X_1 + \cdots + X_n$. Then

$$E(T_n) = n\overline{a}_N \qquad \text{and} \qquad Var(T_n) = n\sigma_a^2\left(1 - \frac{n-1}{N-1}\right) ,$$

and hence for $\overline{X}_n \equiv T_n/n$,

$$E(\overline{X}_n) = \overline{a}_N \qquad \text{and} \qquad Var(\overline{X}_n) = \frac{1}{n}\sigma_a^2\left(1 - \frac{n-1}{N-1}\right) .$$

Here $Cov(X_1, X_2) = -\sigma_a^2/(N-1)$. The "correction factor" $(1-(n-1)/(N-1))$ is due to the covariance no longer being 0. It has the same value no matter what numbers $a_1, \ldots, a_N$ are in the urn.

Proof.

$$0 = Var(X_1 + \cdots + X_N) = N\sigma_a^2 + N(N-1)Cov(X_1, X_2)$$

so $Cov(X_1, X_2) = -\sigma_a^2/(N-1)$. Then

$$\begin{aligned} Var(X_1 + \cdots + X_n) &= n\sigma_a^2 + n(n-1)Cov(X_1, X_2) \\ &= n\sigma_a^2 + n(n-1)\left(-\frac{\sigma_a^2}{N-1}\right) \\ &= n\sigma_a^2\left(1 - \frac{n-1}{N-1}\right). \end{aligned}$$

Example 5. (Variance of Hypergeometric). If the numbers on the balls are $a_1 = \ldots = a_R = 1$, $a_{R+1} = \ldots = a_N = 0$, then $T_n = X_1 + \cdots + X_n =$ the number of red balls in the sample when the urn contains R red balls. Thus $T_n \sim$ Hypergeometric(R, N, n). The results of Example 4 tell us in this special case that

$$E(T_n) = n\overline{a}_N = n\frac{R}{N},$$

and

$$Var(T_n) = \left(1 - \frac{n-1}{N-1}\right)n\sigma_a^2 = \left(1 - \frac{n-1}{N-1}\right)n(R/N)(1 - R/N)$$

since

$$\sigma_a^2 = \frac{1}{N}\sum_{i=1}^{N} a_i^2 - \overline{a}_N^2 = R/N - (R/N)^2 = (R/N)(1 - R/N).$$

Note that these expressions differ only by the finite sampling correction factor $(1 - (n-1)/(N-1))$ from the comparable expressions for the mean and variance of T_n *if the sampling is carried out with replacement*: in that case $T_n \sim$ Binomial$(n, p = R/N)$ and

$$E(T_n) = np = nR/N, \qquad Var(T_n) = np(1-p) = n(R/N)(1 - R/N).$$