15.063: Communicating with Data
Summer 2003

Recitation 3

Probability II

Today's Goal

- Binomial Random Variables (RV)
- Covariance and Correlation
- Sums of RV
- Normal RV

Random Variables

- A random variable assigns a value *(probability)* to each possible outcome of a probabilistic experiment.
- A discrete RV can take only distinct, separate values.
- A continuous RV can take any number.

Binomial Distribution

- **Count** the number of times something happens
- Events have to be **repeated** and **independent**
- Allow us to compute **expectation**, **variance** and **probability of outcomes**
- Described by: **# trials** and **success probability**

Binomial Distribution

- *Example* : flipping a coin 10 times.
  RV number of tails is binomial
  - # trials: 10
  - success probability in each trial: 1/2
- *Question* : Is the number of aces we get in a poker hand a binomial RV?

Binomial Distribution

- *Example* : flipping a coin 10 times.

  X: RV number of tails is binomial(10,1/2)

  - E(**X**) = *np* = 10/2 = 5 (expected # of tails)
  - V(**X**) = *np* (1-*p*) = 10 / 4 = 2.5
  - stdev(**X**) = $\sqrt{np(1-p)} = \sqrt{2.5} = 1.6$
  - p(a single tail) = 5! / 4! x $(.5)^9$ x $(.5)^1$

Taxes

- The probability that your federal tax return will be audited next year is about 0.06 if you have not been audited in the previous three years; this probability increases to 0.12 if you have been audited in the previous three years.

Taxes

(a) Suppose that nine taxpayers are randomly selected, and that none of them have been audited in the past three years.

- What is the probability that exactly one of them will be audited next year?
- What is the probability that more than one of them will be audited next year?
- P(audited) = p = 0.06 and n = 9
- Y = # people audited is Binomial(0.06,9)

$$P(Y=k)=\binom{n}{k}p^k(1-p)^{n-k}=\frac{9!}{(9-k)!k!}0.06^k 0.94^{9-k}$$

Taxes

(a)…Therefore the probability that exactly one is audited is *P(Y=1)* :

$$P(Y=1)=\binom{n}{1}p^1(1-p)^{n-1}=\frac{9!}{8!1!}0.06^1 0.94^8$$

$$=9*0.06*0.94^8$$

$$=0.329167$$

(a) and the probability that more than one is audited is *P(Y>1)* :

$$P(Y>1)=1-P(Y\leq 1)=1-P(Y=0)-P(Y=1)$$

$$\text{we need}: \; P(Y=0)=\frac{9!}{9!0!}0.06^0 0.94^9=0.94^9$$

$$=0.572995$$

$$\Rightarrow P(Y>1)=0.097838$$

Taxes

(b) Suppose that six taxpayers are randomly selected, and that each of them has been audited in the past three years.

- What is the probability that exactly one of them will be audited next year?
- What is the probability that more than one of them will be audited next year?
- P(audited) = p = 0.12 and n = 6
- Y = # people audited is Binomial(0.12,6)

$$P(Y=k)=\binom{n}{k}p^k(1-p)^{n-k}=\frac{6!}{(6-k)!k!}0.12^k 0.88^{6-k}$$

Taxes

(b)…Therefore the probability that exactly one is audited is $P(Y=1)$:

$$P(Y=1)=\binom{n}{1}p^1(1-p)^{n-1}=\frac{6!}{5!1!}0.12^1 0.88^5$$

$$=6*0.12*0.88^5$$

$$=0.379967$$

(b) and the probability that more than one is audited is $P(Y>1)$:

$$P(Y>1)=1-P(Y\leq 1)=1-P(Y=0)-P(Y=1)$$

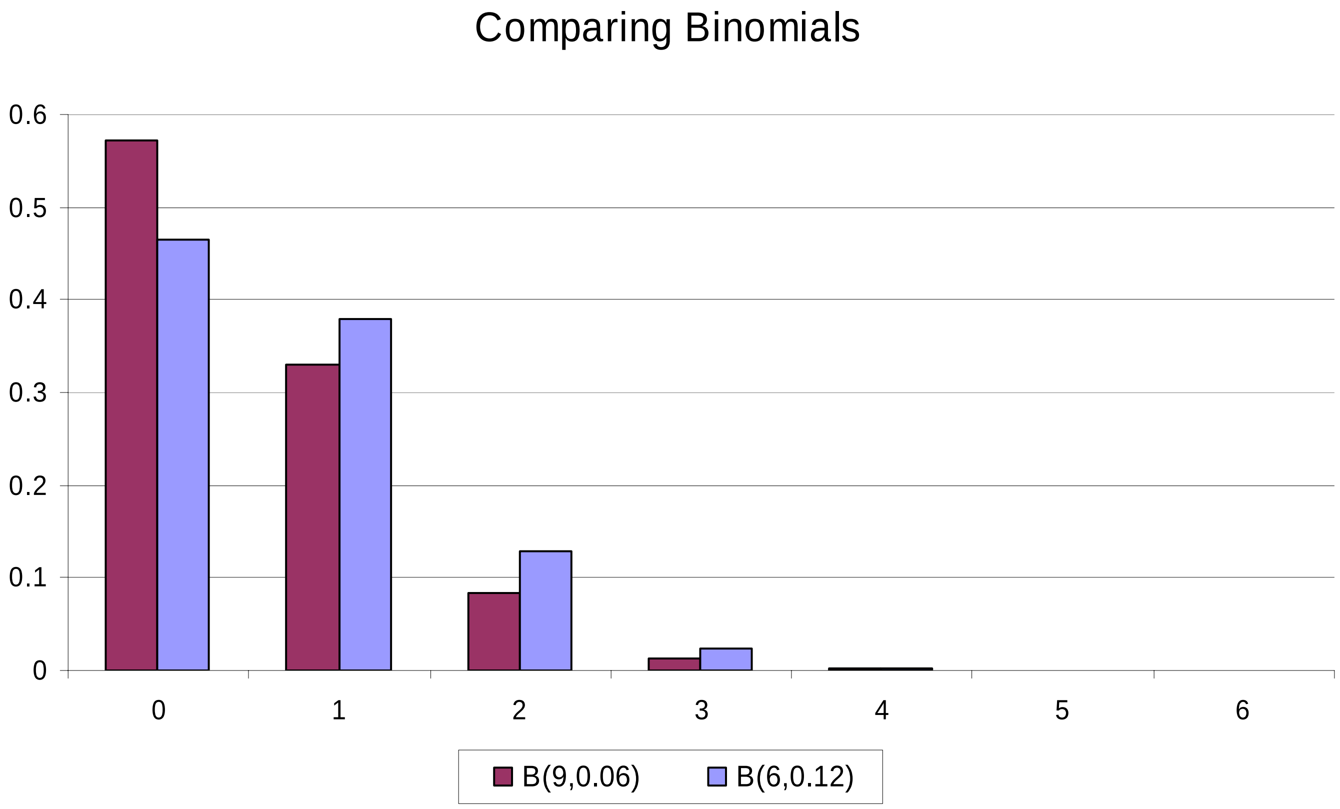
$$\text{we need: } P(Y=0)=\frac{6!}{6!0!}0.12^0 0.88^6=0.88^6$$

$$=0.464404$$

$$\Rightarrow P(Y>1)=0.155629$$

Taxes

- The two binomial distributions we worked with are slightly different. Graphically:



Taxes

(c) If five taxpayers are randomly selected and exactly two of them have been audited in the past three years, what is the probability that none of these taxpayers will be audited by the IRS next year?

We have two binomial distributions $Y \sim B(n1,p1)$ and $Z \sim B(n2,p2)$: where $n1 = 3, p1 = 0.06$ and $n2 = 2, p2 = 0.12$ and we want to compute $P(Y+Z = 0)$:

Since Y can take values *0,1,2,3* and Z can take values *0,1,2,* and they are independent variables:

$$P(Y+Z=0)=P(Y=0 \& Z=0)=P(Y=0)P(Z=0)$$

$$=0.94^3 0.88^2 = 0.643204$$

Covariance and Correlation

- Covariance

$$\text{Cov}(X,Y)=\sum_i P(X=x_i,\ Y=y_i)[(x_i-\mu_x)\ (y_i-\mu_y)]$$

- Correlation

$$\text{CORR}(X,\ Y) = \frac{\text{COV}(X,\ Y)}{\sigma_X \sigma_Y}$$

Sum of two RV

- Mean of sum of two random variables

$$E(aX) = aE(X)$$

$$E(aX + bY) = aE(X) + bE(Y)$$

- Variance of sum of two random variables

$$VAR(aX) = a^2 VAR(X)$$

$$VAR(aX + bY) = a^2 VAR(X) + b^2 VAR(Y) + 2ab\,COV(X, Y)$$

or, equivalently:

$$VAR(aX + bY) = a^2 VAR(X) + b^2 VAR(Y) + 2ab\sigma_X \sigma_Y CORR(X, Y)$$

Stock Portfolio

- A firm is considering a portfolio of stocks. Included in the portfolio are stocks A and B. Let X denote the return from stock A in the following year, and let Y denote the return from stock B.

$E(X) = 0.1$ $\qquad$ $E(Y) = 0.2$

$VAR(X) = 0.0016$ $\qquad$ $VAR(Y) = 0.0036$

$$COV(X,Y) = -0.001$$

Stock Portfolio

(a) What is the expected return of investing 50% in A and 50% in B?

- Let Z be the return: $Z = 0.5\,X + 0.5\,Y$
- $E(Z) = E[0.5\,X + 0.5\,Y]$

  $= 0.5\,E(X) + 0.5\,E(Y)$

  $= (0.5)(0.1) + (0.5)(0.2)$

  $= 0.15$

X is return from Stock A	*Y is return from stock B*
$E(X) = 0.1$	$E(Y) = 0.2$
$VAR(X) = 0.0016$	$VAR(Y) = 0.0036$

$COV(X,Y) = -0.001$

Stock Portfolio

(b) What is the standard deviation of this return?

- $VAR(a X + b Y) = a^2 VAR(X) + b^2 VAR(Y) + 2ab\, COV(X,Y)$
- $VAR(Z) = VAR(0.5 X + 0.5 Y)$

  $= (0.5)^2(0.0016) + (0.5)^2(0.0036)$

  $+ (2)(0.5)(0.5)(-0.001) = 0.0008$
- $\sigma_z = \sqrt{} VAR(Z) = \sqrt{} 0.008 = 0.0283$

X is return from Stock A	*Y is return from stock B*
$E(X) = 0.1$	$E(Y) = 0.2$
$VAR(X) = 0.0016$	$VAR(Y) = 0.0036$

$COV(X,Y) = -0.001$

Stock Portfolio

- (c) What is the correlation between X and Y?
  - $COV(X,Y) = CORR(X,Y)\ \sigma_X\ \sigma_Y$
  - $CORR(X,Y) = (-0.001)/(\sigma_X\ \sigma_Y)$

$$= (-0.001)/[(\sqrt{VAR(X)})(\sqrt{VAR(Y)}]$$
$$= (-0.001)/[(\sqrt{0.0016})(\sqrt{0.0036})]$$
$$= (-0.001)/[(0.04)(0.06)]$$
$$= -0.4167$$

X is return from Stock A	*Y is return from stock B*
E(X) = 0.1	*E(Y) = 0.2*
VAR(X) = 0.0016	*VAR(Y) = 0.0036*

COV(X,Y) = -0.001

Stock Portfolio

(d) What should be the composition of the portfolio if the firm wants an expected return of 18%?

- Let p be the % of Stock A in portfolio.
- $E[pX+(1-p)Y] = pE(X)+(1-p)E(Y)$
  $= p(0.1) + (1-p)(0.2)$
  $= 0.1p + 0.2 - 0.2p = -0.1p + 0.2$
- Setting $-0.1p + 0.2 = 0.18$, we get $p = 0.2$

Portfolio: 20% of Stock A and 80% of Stock B

X is return from Stock A	*Y is return from stock B*
$E(X) = 0.1$	$E(Y) = 0.2$
$VAR(X) = 0.0016$	$VAR(Y) = 0.0036$
$COV(X,Y) = -0.001$	

Stock Portfolio

(e) What is the standard deviation of the return?

- $VAR(aX + bY) = a^2\, VAR(X) + b^2\, VAR(Y) + 2ab\, COV(X,Y)$
- $VAR(Z') = VAR(0.2\, X + 0.8\, Y)$

  $= (0.2)^2(0.0016) + (0.8)^2(0.0036) + (2)(0.2)(0.8)(-0.001)$

  $= 0.002048$
- $\sigma_{Z'} = \sqrt{VAR(Z')} = \sqrt{0.002048} = 0.0453$

X is return from Stock A	*Y is return from stock B*
E(X) = 0.1	*E(Y) = 0.2*
VAR(X) = 0.0016	*VAR(Y) = 0.0036*
COV(X,Y) = -0.001	

Standard Normal Distribution

- *Standard normal RV* : bell shaped distribution
  Center in 0 and Std. Deviation =1
- Widely used in practice to model uncertainty
- *Denoted* : $Z \sim N(0,1)$
- *Cumulative Distribution*: $F(z)=P(Z \leq z)$ for $Z \sim N(0,1)$
  F(z) can be found in the Normal Table or computed using Excel.

Normal Distribution

- *Any normal RV* : bell shaped distribution
  Center in μ and Std. Deviation= σ
- *Denoted* : N(μ, σ)
- If X is any normal RV (i.e., $X \sim N(\mu, \sigma)$),
  then $Z = (X - \mu)/\sigma$ is a standard normal RV
- This enables us to obtain values for any $X \sim N(\mu, \sigma)$
- Example: $X \sim N(2,3)$
  - What is the probability that $X < 4$?

Sum of Normal Distributions

- If X and Y are normally distributed, the sum $X + Y$ is also normally distributed.
- Its mean and variance are computed with the ordinary formulas.
- Example: $X \sim$ N(2,3) and $Y \sim$ N(1,4)
  - $X + Y \sim N(2+1, \sqrt{3^2 + 4^2})$

Normal Distribution

- The weekly price <u>change</u> of a share of stock X is normally distributed with mean $0.05P$ and variance 1, where P is the price at the beginning of the week.

(a) If a share of stock X costs \$24 at the beginning of a week, what is the probability the stock goes up that week?

(b) Given that the stock goes up that week, what is the probability it reaches \$27?

The End.