Introduction to Computational Finance and Financial Econometrics

Probability Review - Part 2

Eric Zivot

Spring 2015

Outline

Example

Example: Two discrete rv's X and Y

Bivariate pdf

		Y		
	%	0	1	Pr(X)
	0	1/8	0	1/8
X	1	2/8	1/8	3/8
	2	1/8	2/8	3/8
	3	0	1/8	1/8
	Pr(Y)	4/8	4/8	1

$$p(x, y) = \Pr(X = x, Y = y) = \text{values in table}$$

$$e.g.,\ p(0, 0) = \Pr(X = 0, Y = 0) = \ 1/8$$

Properties of joint pdf $p(x, y)$

$$S_{XY} = \{(0,0),\ (0,1),\ (1,0),\ (1,1),$$
$$(2,0),\ (2,1),\ (3,0),\ (3,1)\}$$

$$p(x,y) \geq 0 \text{ for } x, y \in S_{XY}$$

$$\sum_{x,y \in S_{XY}} p(x,y) = 1$$

Outline

Marginal pdfs

$$p(x) = \Pr(X = x) = \sum_{y \in S_Y} p(x, y)$$

$$= \text{sum over columns in joint table}$$

$$p(y) = \Pr(Y = y) = \sum_{x \in S_X} p(x, y)$$

$$= \text{sum over rows in joint table}$$

Outline

Conditional Probability

Suppose we know $Y = 0$. How does this knowledge affect the probability that $X = 0, 1, 2,$ or 3? The answer involves conditional probability. **Example:**

$$\Pr(X = 0|Y = 0) = \frac{\Pr(X = 0, Y = 0)}{\Pr(Y = 0)}$$

$$= \frac{\text{joint probability}}{\text{marginal probability}} = \frac{1/8}{4/8} = 1/4$$

Remark:

$$\Pr(X = 0|Y = 0) = 1/4 \neq \Pr(X = 0) = 1/8$$

$$\implies X \text{ depends on } Y$$

The marginal probability, $\Pr(X = 0)$, ignores information about Y.

Definition - Conditional Probability

- The conditional pdf of X given $Y = y$ is, for all $x \in S_X$,

$$p(x|y) = \Pr(X = x|Y = y) = \frac{\Pr(X = x, Y = y)}{\Pr(Y = y)}$$

- The conditional pdf of Y given $X = x$ is, for all values of $y \in S_Y$

$$p(y|x) = \Pr(Y = y|X = x) = \frac{\Pr(X = x, Y = y)}{\Pr(X = x)}$$

Conditional Mean and Variance

$$\mu_{X|Y=y} = E[X|Y=y] = \sum_{x \in S_X} x \cdot \Pr(X=x|Y=y),$$

$$\mu_{Y|X=x} = E[Y|X=x] = \sum_{y \in S_Y} y \cdot \Pr(Y=y|X=x).$$

$$\sigma^2_{X|Y=y} = \mathrm{var}(X|Y=y) = \sum_{x \in S_X} (x - \mu_{X|Y=y})^2 \cdot \Pr(X=x|Y=y),$$

$$\sigma^2_{Y|X=x} = \mathrm{var}(Y|X=x) = \sum_{y \in S_Y} (y - \mu_{Y|X=x})^2 \cdot \Pr(Y=y|X=x).$$

Example

$$E[X] = 0 \cdot 1/8 + 1 \cdot 3/8 + 2 \cdot 3/8 + 3 \cdot 1/8 = 3/2$$

$$E[X|Y = 0] = 0 \cdot 1/4 + 1 \cdot 1/2 + 2 \cdot 1/4 + 3 \cdot 0 = 1,$$

$$E[X|Y = 1] = 0 \cdot 0 + 1 \cdot 1/4 + 2 \cdot 1/2 + 3 \cdot 1/4 = 2,$$

$$\text{var}(X) = (0 - 3/2)^2 \cdot 1/8 + (1 - 3/2)^2 \cdot 3/8$$

$$+(2 - 3/2)^2 \cdot 3/8 + (3 - 3/2)^2 \cdot 1/8 = 3/4,$$

$$\text{var}(X|Y = 0) = (0 - 1)^2 \cdot 1/4 + (1 - 1)^2 \cdot 1/2$$

$$+(2 - 1)^2 \cdot 1/2 + (3 - 1)^2 \cdot 0 = 1/2,$$

$$\text{var}(X|Y = 1) = (0 - 2)^2 \cdot 0 + (1 - 2)^2 \cdot 1/4$$

$$+(2 - 2)^2 \cdot 1/2 + (3 - 2)^2 \cdot 1/4 = 1/2.$$

Independence

Let X and Y be discrete rvs with pdfs $p(x)$, $p(y)$, sample spaces S_X, S_Y and joint pdf $p(x, y)$. Then X and Y are independent rv's if and only if:

$$p(x, y) = p(x) \cdot p(y)$$

for all values of $x \in S_X$ and $y \in S_Y$

Result: If X and Y are independent rv's then,

$$p(x|y) = p(x) \text{ for all } x \in S_X,\ y \in S_Y$$

$$p(y|x) = p(y) \text{ for all } x \in S_X,\ y \in S_Y$$

Intuition:

- Knowledge of X does not influence probabilities associated with Y
- Knowledge of Y does not influence probablities associated with X

Outline

Bivariate Distributions - Continuous rv's

The joint pdf of X and Y is a non-negative function $f(x, y)$ such that:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$$

Let $[x_1, x_2]$ and $[y_1, y_2]$ be intervals on the real line. Then,

$$\Pr(x_1 \leq X \leq x_2, y_1 \leq Y \leq y_2)$$

$$= \int_{x_1}^{x_2} \int_{y_1}^{y_2} f(x, y) dx dy$$

$$= \text{volume under probability surface}$$

over the intersection of the intervals

$[x_1, x_2]$ and $[y_1, y_2]$

Outline

1. Bivariate Probability Distribution - Discrete rv's
2. Bivariate Probability Distribution - Continuous rv's
   - Marginal and conditional distributions
3. Independence
4. Covariance and Correlation

Marginal and conditional distributions

The marginal pdf of X is found by integrating y out of the joint pdf $f(x,y)$ and the marginal pdf of Y is found by integrating x out of the joint pdf:

$$f(x) = \int_{-\infty}^{\infty} f(x,y)dy,$$

$$f(y) = \int_{-\infty}^{\infty} f(x,y)dx.$$

The conditional pdf of X given that $Y = y$, denoted $f(x|y)$, is computed as,

$$f(x|y) = \frac{f(x,y)}{f(y)},$$

and the conditional pdf of Y given that $X = x$ is computed as,

$$f(y|x) = \frac{f(x,y)}{f(x)}.$$

Conditional mean and variance

The conditional means are computed as:

$$\mu_{X|Y=y} = E[X|Y=y] = \int x \cdot p(x|y)dx,$$

$$\mu_{Y|X=x} = E[Y|X=x] = \int y \cdot p(y|x)dy$$

and the conditional variances are computed as,

$$\sigma^2_{X|Y=y} = \mathrm{var}(X|Y=y) = \int (x - \mu_{X|Y=y})^2 p(x|y)dx,$$

$$\sigma^2_{Y|X=x} = \mathrm{var}(Y|X=x) = \int (y - \mu_{Y|X=x})^2 p(y|x)dy.$$

Outline

Independence

Let X and Y be continuous random variables. X and Y are independent iff:

$$f(x|y) = f(x), \text{ for } -\infty < x, y < \infty,$$

$$f(y|x) = f(y), \text{ for } -\infty < x, y < \infty.$$

Result: Let X and Y be continuous random variables. X and Y are independent iff:

$$f(x, y) = f(x)f(y)$$

The result in the above proposition is extremely useful in practice because it gives us an easy way to compute the joint pdf for two independent random variables: we simple compute the product of the marginal distributions.

Example

Example: Bivariate standard normal distribution Let $X \sim N(0,1)$, $Y \sim N(0,1)$ and let X and Y be independent. Then,

$$f(x,y) = f(x)f(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}$$

$$= \frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}.$$

To find $\Pr(-1 < X < 1, -1 < Y < 1)$ we must solve,

$$\int_{-1}^{1}\int_{-1}^{1}\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}dxdy$$

which, unfortunately, does not have an analytical solution. Numerical approximation methods are required to evaluate the above integral. See R package mvtnorm.

Independence cont.

Result: If the random variables X and Y (discrete or continuous) are independent then the random variables $g(X)$ and $h(Y)$ are independent for any functions $g(\cdot)$ and $h(\cdot)$.

For example, if X and Y are independent then X^2 and Y^2 are also independent.

Outline

Covariance and Correlation - Measuring linear dependence between two rv's

Covariance: Measures direction but not strength of linear relationship between 2 rv's

$$
\begin{aligned}
\sigma_{XY} &= E[(X-\mu_X)(Y-\mu_Y)] \\
&= \sum_{x.y \in S_{XY}} (x-\mu_X)(y-\mu_Y) \cdot p(x,y) \text{ (discrete)} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y) f(x,y) dxdy \text{ (cts)}
\end{aligned}
$$

Example

Example: For the data in Table 2, we have

$$\sigma_{XY} = \text{Cov}(X,Y) = (0-3/2)(0-1/2)\cdot 1/8$$

$$+(0-3/2)(1-1/2)\cdot 0 + \cdots$$

$$+(3-3/2)(1-1/2)\cdot 1/8 = 1/4$$

Properties of Covariance

$$\text{Cov}(X, Y) = \text{Cov}(Y, X)$$

$$\text{Cov}(aX, bY) = a \cdot b \cdot \text{Cov}(X, Y) = a \cdot b \cdot \sigma_{XY}$$

$$\text{Cov}(X, X) = \text{Var}(X)$$

$$X, Y \text{ independent} \implies \text{Cov}(X, Y) = 0$$

$$\text{Cov}(X, Y) = 0 \nRightarrow X \text{ and } Y \text{ are independent}$$

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$$

Correlation

Correlation: Measures direction and strength of linear relationship between 2 rv's

$$\rho_{XY} = \mathrm{Cor}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\mathrm{SD}(X) \cdot \mathrm{SD}(Y)}$$

$$= \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y} = \text{ scaled covariance}$$

Example

Example: For the Data in Table 2

$$\rho_{XY} = \mathrm{Cor}(X,Y) = \frac{1/4}{\sqrt{(3/4)\cdot(1/2)}} = 0.577$$

Properties of Correlation

$$-1 \leq \rho_{XY} \leq 1$$

$$\rho_{XY} = 1 \text{ if } Y = aX + b \text{ and } a > 0$$

$$\rho_{XY} = -1 \text{ if } Y = aX + b \text{ and } a < 0$$

$$\rho_{XY} = 0 \text{ if and only if } \sigma_{XY} = 0$$

$$\rho_{XY} = 0 \nRightarrow \; X \text{ and } Y \text{ are independent in general}$$

$$\rho_{XY} = 0 \implies \text{independence if } X \text{ and } Y \text{ are normal}$$

Outline

1. Bivariate Probability Distribution - Discrete rv's
2. Bivariate Probability Distribution - Continuous rv's
3. Independence
4. Covariance and Correlation
   - Bivariate normal distribution
   - Linear Combination of rv's

Bivariate normal distribution

Let X and Y be distributed bivariate normal. The joint pdf is given by:

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \times$$

$$\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right\}$$

where $E[X] = \mu_X$, $E[Y] = \mu_Y$, $\mathrm{SD}(X) = \sigma_X$, $\mathrm{SD}(Y) = \sigma_Y$, and $\rho = \mathrm{cor}(X,Y)$.

Outline

Linear Combination of 2 rv's

Let X and Y be rv's. Define a new rv Z that is a linear combination of X and Y:

$$Z = aX + bY$$

where a and b are constants. Then,

$$\begin{aligned}\mu_Z &= E[Z] = E[aX + bY] \\ &= aE[X] + bE[Y] \\ &= a \cdot \mu_X + b \cdot \mu_Y\end{aligned}$$

Linear Combination of 2 rv's cont.

and,

$$\begin{aligned}\sigma_Z^2 &= \text{Var}(Z) = \text{Var}(a \cdot X + b \cdot Y) \\ &= a^2\text{Var}(X) + b^2\text{Var}(Y) + 2a \cdot b \cdot \text{Cov}(X, Y) \\ &= a^2\sigma_X^2 + b^2\sigma_Y^2 + 2a \cdot b \cdot \sigma_{XY}\end{aligned}$$

If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ then $Z \sim N(\mu_Z, \sigma_Z^2)$.

Example

Example: Portfolio returns

- R_A = return on asset A with $E[R_A] = \mu_A$ and $\text{Var}(R_A) = \sigma_A^2$
- R_B = return on asset B with $E[R_B] = \mu_B$ and $\text{Var}(R_B) = \sigma_B^2$
- $\text{Cov}(R_A, R_B) = \sigma_{AB}$ and $\text{Cor}(R_A,\ R_B) = \rho_{AB} = \frac{\sigma_{AB}}{\sigma_A \cdot \sigma_B}$

Portfolio:

- x_A = share of wealth invested in asset A, x_B = share of wealth invested in asset B
- $x_A + x_B = 1$ (exhaust all wealth in 2 assets)
- $R_P = x_A \cdot R_A + x_B \cdot R_B$ = portfolio return

Problem

Portfolio Problem: How much wealth should be invested in assets A and B?

Portfolio expected return (gain from investing):

$$\begin{aligned} E[R_P] &= \mu_P = E[x_A \cdot R_A + x_B \cdot R_B] \\ &= x_A E[R_A] + x_B E[R_B] \\ &= x_A \mu_A + x_B \mu_B \end{aligned}$$

Problem cont.

Portfolio variance (risk from investing):

$$\begin{aligned}\mathrm{Var}(R_P) &= \sigma_P^2 = \mathrm{Var}(x_A R_A + x_B R_B) \\ &= x_A^2 \mathrm{Var}(R_A) + x_B^2 \mathrm{Var}(R_B) + \\ & 2 \cdot x_A \cdot x_B \cdot \mathrm{Cov}(R_A,\ R_B) \\ &= x_A^2 \sigma_A^2 + x_B^2 \sigma_B^2 + 2 x_A x_B \sigma_{AB} \\ \mathrm{SD}(R_P) &= \sqrt{\mathrm{Var}(R_P)} = \sigma_P \\ &= \left(x_A^2 \sigma_A^2 + x_B^2 \sigma_B^2 + 2 x_A x_B \sigma_{AB} \right)^{1/2}\end{aligned}$$

Linear Combination of N rv's

Let $X_1,\ X_2, \cdots, X_N$ be rvs and let $a_1, a_2, \ldots, a_N$ be constants. Define,

$$Z = a_1 X_1 + a_2 X_2 + \cdots + a_N X_N = \sum_{i=1}^{N} a_i X_i$$

Then,

$$\mu_Z = E[Z] = a_1 E[X_1] + a_2 E[X_2] + \cdots + a_N E[X_N]$$

$$= \sum_{i=1}^{N} a_i E[X_i] = \sum_{i=1}^{N} a_i \mu_i$$

Linear Combination of N rv's cont.

For the variance,

$$\sigma_Z^2 = \text{Var}(Z) = a_1^2\text{Var}(X_1) + \cdots + a_N^2\text{Var}(X_N)$$

$$+ 2a_1a_2\text{Cov}(X_1,\ X_2) + 2a_1a_3\text{Cov}(X_1,\ X_3) + \cdots$$

$$+ 2a_2a_3\text{Cov}(X_2,\ X_3) + 2a_2a_4\text{Cov}(X_2,\ X_4) + \cdots$$

$$+ 2a_{N-1}a_N\text{Cov}(X_{N-1}, X_N)$$

Note: N variance terms and $N(N-1) = N^2 - N$ covariance terms. If $N = 100$, there are $100 \times 99 = 9900$ covariance terms!

Result: If $X_1,\ X_2, \cdots, X_N$ are each normally distributed random variables then,

$$Z = \sum_{i=1}^{N} a_i X_i \sim N(\mu_Z, \sigma_Z^2)$$

Example: Portfolio variance with three assets

- R_A, R_B, R_C are simple returns on assets A, B and C
- x_A, x_B, x_C are portfolio shares such that $x_A + x_B + x_C = 1$
- $R_p = x_A R_A + x_B R_B + x_C R_C$

Portfolio variance,

$$\sigma_P^2 = x_A^2\sigma_A^2 + x_B^2\sigma_B^2 + x_C^2\sigma_C^2$$
$$+ 2x_A x_B \sigma_{AB} + 2x_A x_C \sigma_{AC} + 2x_B x_C \sigma_{BC}$$

Note: Portfolio variance calculation may be simplified using matrix layout,

	x_A	x_B	x_C
x_A	σ_A^2	σ_{AB}	σ_{AC}
x_B	σ_{AB}	σ_B^2	σ_{BC}
x_C	σ_{AC}	σ_{BC}	σ_C^2

Example: Multi-period continuously compounded returns and the square-root-of-time rule

$$r_t = \ln(1 + R_t) = \text{ monthly cc return}$$

$$r_t \sim N(\mu,\ \sigma^2) \text{ for all } t$$

$$\text{Cov}(r_t, r_s) = 0 \text{ for all } t \neq s$$

Annual return,

$$r_t(12) = \sum_{j=0}^{11} r_{t-j}$$

$$= r_t + r_{t-1} + \cdots + r_{t-11}$$

Example cont.

Then,

$$
\begin{aligned}
E[r_t(12)] &= \sum_{j=0}^{11} E[r_{t-j}] \\
&= \sum_{j=0}^{11} \mu \ (E[r_t] = \mu \text{ for all } t) \\
&= 12\mu \ (\mu = \text{mean of monthly return})
\end{aligned}
$$

Example cont.

And,

$$\text{Var}(r_t(12)) = \text{Var}\left(\sum_{j=0}^{11} r_{t-j}\right)$$

$$= \sum_{j=0}^{11} \text{Var}(r_{t-j}) = \sum_{j=0}^{11} \sigma^2$$

$$= 12 \cdot \sigma^2 \ (\sigma^2 = \text{monthly variance})$$

$$\text{SD}(r_t(12)) = \sqrt{12} \cdot \sigma \ (\text{square root of time rule})$$

Then,

$$r_t(12) \sim N(12\mu, 12\sigma^2)$$

Example cont.

For example, suppose:

$$r_t \sim N(0.01, (0.10)^2)$$

Then,

$$E[r_t(12)] = 12 \times (0.01) = 0.12$$

$$\text{Var}(r_t(12)) = 12 \times (0.10)^2 = 0.12$$

$$\text{SD}(r_t(12)) = \sqrt{0.12} = 0.346$$

$$r_t(12) \sim N(0.12, (0.346)^2)$$

Example cont.

And,

$$\begin{aligned}(q_\alpha^r)^A &= 12 \times \mu + \sqrt{12} \times \sigma \times z_\alpha \\ &= 0.12 + 0.346 \times z_\alpha \\ (q_\alpha^R)^A &= e^{(q_\alpha^r)^A} - 1 = e^{0.12 + 0.346 \times z_\alpha} - 1\end{aligned}$$

Covariance between two linear combinations of random variables

Consider two linear combinations of two random variables:

$$X = X_1 + X_2$$

$$Y = Y_1 + Y_2$$

Then,

$$\begin{aligned}\text{cov}(X, Y) &= \text{cov}(X_1 + X_2, Y_1 + Y_2) \\ &= \text{cov}(X_1, Y_1) + \text{cov}(X_1, Y_2) \\ &+ \text{cov}(X_2, Y_1) + \text{cov}(X_2, Y_2)\end{aligned}$$

The result generalizes to linear combinations of N random variables in the obvious way.

faculty.washington.edu/ezivot/