Some Probability for Math of Finance I

1. A random variable $X$ is \textit{discrete} if it can only take on discrete values.

Example: Suppose $x_1 < x_2 < .. < x_m$ are constants. Let

$$X = \begin{cases} 
  x_1 & \text{with probability } p_1 \\
  x_2 & \text{with probability } p_2 \\
  \vdots & \\
  x_m & \text{with probability } p_m 
\end{cases}$$

Then for $i = 1, 2, .., m$ the probability that $X$ equals $x_i$ is $P[X = x_i] = p_i$.

A random variable $X$ is \textit{continuous} if it can take on any value over a continuous range. Example: $X$ takes on values $-\infty < X < \infty$ and the probability that $X$ falls in subinterval $[a, b]$ is

$$P[a \leq X \leq b] = \int_{a}^{b} p(x)dx$$

for some probability density $p$.

2. A function $g$ is given. If $X$ discrete, the expected value $E[g(X)]$ is the sum of the values $g(X)$ can take, weighted by the probabilities of the values. For the example above,

$$E[g(X)] = g(x_1)p_1 + g(x_2)p_2 + .. + g(x_m)p_m$$

If $X$ continuous, the expected value $E[g(X)]$ is the integral of the values $g(X)$ can take, weighted by the probability density $p$. For the example above,

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)p(x)dx$$

If instead of $X$ the symbol $x$ is used for the random variable, the formulas for the expected values become

$$E[g(x)] = g(x_1)p_1 + g(x_2)p_2 + .. + g(x_m)p_m \text{ and}$$

$$E[g(x)] = \int_{-\infty}^{\infty} g(\xi)p(\xi)d\xi, \ \ \xi \text{ pronounced “ksi” with the “ks” like the \textit{x} in fox}$$
3. If $X$ discrete, the mean and variance of $X$ are

$$\bar{X} = E[X] = \sum_{i=1}^{m} x_i p_i$$

$$\text{var}(X) = E[(X - \bar{X})^2] = \sum_{i=1}^{m} (x_i - \bar{X})^2 p_i$$

If $X$ continuous, the mean and variance of $X$ are

$$E[X] = \int_{-\infty}^{\infty} x p(x) \, dx$$

$$\text{var}(X) = E[(X - \bar{X})^2] = \int_{-\infty}^{\infty} (x - \bar{X})^2 p(x) \, dx$$

4. Given a function $g$, the mean of $g(X)$ is

$$\bar{g} = E[g(X)] = \begin{cases} \sum_{i=1}^{m} g(x_i) p_i \\ \int_{-\infty}^{\infty} g(x) p(x) \, dx \end{cases} \quad \text{if } X \text{ discrete}$$

and the variance is

$$\text{var}(g(X)) = E[(g(X) - \bar{g})^2] = \begin{cases} \sum_{i=1}^{m} (g(x_i) - \bar{g})^2 p_i \\ \int_{-\infty}^{\infty} (g(x) - \bar{g})^2 p(x) \, dx \end{cases} \quad \text{if } X \text{ continuous}$$

5. Given functions $g$ and $h$, the covariance of $g(X)$ and $h(X)$ is

$$\text{cov}(g(X), h(X)) = E[(g(X) - \bar{g})(h(X) - \bar{h})]$$

where $\bar{g} = E[g(X)]$, $\bar{h} = E[h(X)]$ are the means.

6. If $X$ and $Y$ are discrete random variables, so is the pair $(X, Y)$.

Example: $X$ takes on values $x_1, x_2, \ldots, x_m$ and $Y$ takes on values $y_1, y_2, \ldots, y_n$. Then

$$(X, Y) = \begin{cases} (x_1, y_1) \text{ with probability } p_{1,1} \\ (x_1, y_2) \text{ with probability } p_{1,2} \\ \vdots \\ (x_i, y_j) \text{ with probability } p_{i,j} \\ \vdots \\ (x_m, y_n) \text{ with probability } p_{m,n} \end{cases}$$
where the probabilities \( p_{i,j} \) sum to 1.

For functions \( g(x, y) \), the expected value \( g(X, Y) \) is

\[
E[g(X, Y)] = \sum_{i=1}^{m} \sum_{j=1}^{n} g(x_i, y_j) p_{i,j}
\]

7. If \( X \) and \( Y \) are continuous random variables, so is the pair \((X, Y)\).

Example: Suppose \(-\infty < X < \infty \) and \(-\infty < Y < \infty \). The probability that the pair \((X, Y)\) falls in any subrectangle \([a, b] \times [c, d]\) is given by the integral

\[
\int_{c}^{d} \int_{a}^{b} p(x, y) \, dx \, dy
\]

where \( p(x, y) \) is the probability density function.

For any function \( g(x, y) \), the expected value of \( g(X, Y) \) is

\[
E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) p(x, y) \, dx \, dy
\]

8. If \( X, Y \) are both random variables as in the discrete example,

\[
\text{cov}[X, Y] = \sum_{i=1}^{m} \sum_{j=1}^{n} (x_i - \bar{X})(y_j - \bar{Y})p_{i,j}
\]

If \( X, Y \) are both random variables as in the continuous example.

\[
\text{cov}[X, Y] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{X})(y - \bar{Y}) p(x, y) \, dx \, dy
\]

In the discrete example, if \( p_{i,j} = p_i q_j \) for some probabilities \( p_i, q_j \), then \( \text{cov}[X, Y] = 0 \). In the continuous example, if the function \( p(x, y) \) is the product of a function of \( x \) times a function of \( y \), then \( \text{cov}[X, Y] = 0 \).

9. If \( r_1, r_2 \) and \( r_3 \) are three continuous random variables which each take on real values \(-\infty < r_i < \infty \) and \( p(x, y, z) \) is the probability density function, then for general \( g(x, y, z) \), the expected value of \( g(r_1, r_2, r_3) \) is

\[
E[g(r_1, r_2, r_3)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y, z) p(x, y, z) \, dx \, dy \, dz
\]
Choosing $g = r_i$, $i = 1, 2, 3$ gives formulas for the mean values
\[ \bar{r}_i = E[r_i], \quad i = 1, 2, 3 \]

Then choosing $g = (r_i - \bar{r}_i)(r_j - \bar{r}_j)$ gives formulas for the covariances
\[ \text{cov}(r_i, r_j) = E[(r_i - \bar{r}_i)(r_j - \bar{r}_j)], \quad i, j = 1, 2, 3 \]

The $3 \times 3$ matrix with elements $\{\text{cov}(r_i, r_j)\}$ is called the covariance matrix. This matrix is written as
\[
\begin{pmatrix}
\text{cov}(r_1, r_1) & \text{cov}(r_1, r_2) & \text{cov}(r_1, r_3) \\
\text{cov}(r_2, r_1) & \text{cov}(r_2, r_2) & \text{cov}(r_2, r_3) \\
\text{cov}(r_3, r_1) & \text{cov}(r_3, r_2) & \text{cov}(r_3, r_3)
\end{pmatrix} =
\begin{pmatrix}
\sigma_{1,1} & \sigma_{1,2} & \sigma_{1,3} \\
\sigma_{2,1} & \sigma_{2,2} & \sigma_{2,3} \\
\sigma_{3,1} & \sigma_{3,2} & \sigma_{3,3}
\end{pmatrix}
\]

in terms of symbols $\sigma_{i,j} = \text{cov}(r_i, r_j)$. Note that for $j = i$, $\sigma_{i,i} = \text{cov}(r_i, r_i) = E[(r_i - \bar{r}_i)(r_i - \bar{r}_i)] = \text{var}(r_i) = \sigma_i^2$, so $\sigma_{1,1} = \sigma_1^2$, $\sigma_{2,2} = \sigma_2^2$, $\sigma_{3,3} = \sigma_3^2$. Also, $\sigma_{i,j} = \text{cov}(r_i, r_j) = \text{cov}(r_j, r_i) = \sigma_{j,i}$, so $\sigma_{2,1} = \sigma_{1,2}$, $\sigma_{3,1} = \sigma_{1,3}$ and $\sigma_{3,2} = \sigma_{2,3}$. The matrix is symmetric and may be written as
\[
\begin{pmatrix}
\sigma_1^2 & \sigma_{1,2} & \sigma_{1,3} \\
\sigma_{1,2} & \sigma_2^2 & \sigma_{2,3} \\
\sigma_{1,3} & \sigma_{2,3} & \sigma_3^2
\end{pmatrix} =
\begin{pmatrix}
\sigma_1^2 & \sigma_{1,2} & \sigma_{1,3} \\
0 & \sigma_2^2 & \sigma_{2,3} \\
0 & 0 & \sigma_3^2
\end{pmatrix}
\]

where the “.” entries in the lower triangle are understood to be copies of the corresponding entries in the upper triangle.

10. If $r_1, r_2, \ldots, r_N$ are $N$ continuous random variables which each take on real values $-\infty < r_i < \infty$ and $p(x_1, x_2, \ldots, x_N)$ is the probability density function, then for given function $g$, $E[g(r_1, r_2, \ldots, r_N)]$ is the repeated integral
\[
\int_{x_N = -\infty}^{\infty} \cdots \int_{x_2 = -\infty}^{\infty} \int_{x_1 = -\infty}^{\infty} g(x_1, x_2, \ldots, x_N) p(x_1, x_2, \ldots, x_N) \, dx_1 \, dx_2 \ldots \, dx_N
\]

The choices $g = x_1$, $g = x_2$, ..., $g = x_N$ give formulas for the mean values
\[ \bar{r}_i = E[r_i], \quad i = 1, 2, \ldots, N \]

and then the choices $g = (x_i - \bar{r}_i)(x_j - \bar{r}_j)$ give the covariances
\[ \text{cov}(r_i, r_j) = E[(r_i - \bar{r}_i)(r_j - \bar{r}_j)], \quad i, j = 1, 2, \ldots, N \]
The covariance matrix is symmetric, and may be written as

\[
\begin{pmatrix}
\sigma^2_1 & \sigma_{1,2} & \sigma_{1,3} & \ldots & \sigma_{1,N} \\
\sigma_{1,2} & \sigma^2_2 & \sigma_{2,3} & \ldots & \sigma_{2,N} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\sigma_{1,N} & \sigma_{2,N} & \sigma_{3,N} & \ldots & \sigma^2_N
\end{pmatrix}
\]

in terms of symbols \( \sigma_{i,j} = \text{cov}(r_i, r_j), \ i, j = 1, \ldots, N \) and \( \sigma_i = \sqrt{\text{var}(r_i)}, i = 1, \ldots, N \).

11. Problems in portfolio construction model yearly return rates for the different assets as random variables \( r_1, \ldots, r_N \). Information about these variables will be given, such as the expected values \( \bar{r}_i \), variances \( \text{var}(r_i) = \sigma^2_i \) and covariances \( \text{cov}(r_i, r_j) = \sigma_{i,j} \) or correlations \( \rho_{i,j} \). Note that

\[
\rho_{i,j} = \frac{\text{cov}(r_i, r_j)}{\sqrt{\text{var}(r_i)} \sqrt{\text{var}(r_j)}}, \quad \text{cov}(r_i, r_j) = \rho_{i,j} \sigma_i \sigma_j \quad i, j = 1, \ldots, N
\]

so it is easy to convert between covariances and correlations.

The value of “beta” might be given instead of a covariance or correlation. For an asset with return rate \( r_a \) relative to a benchmark with return rate \( r_b \),

\[
\beta = \frac{\text{cov}(r_a, r_b)}{\text{var}(r_b)}, \quad \text{cov}(r_a, r_b) = \beta \text{var}(r_b)
\]

12. Since the expected value of an expression can in principle be found by multiplying the expression by \( p \) and then integrating, the expected value \( E[] \) acts linearly. For example, for any constants \( c_1 \) and \( c_2 \) (constants do not depend on the random variables ), \( E[c_1 r_1 + c_2 r_2] = c_1 E[r_1] + c_2 E[r_2] \).

13. Since the covariance is defined in terms of and expected value, the covariance acts linearly with respect to the first of its arguments.

\[
\text{cov}(c_1 r_1 + c_2 r_2, r_3) = E[(c_1 (r_1 - \bar{r}_1) + c_2 (r_2 - \bar{r}_2))(r_3 - \bar{r}_3)] \\
= c_1 E[(r_1 - \bar{r}_1)(r_3 - \bar{r}_3)] + c_2 E[(r_2 - \bar{r}_2)(r_3 - \bar{r}_3)] = c_1 \text{cov}(r_1, r_3) + c_2 \text{cov}(r_2, r_3)
\]

for any constants \( c_1 \) and \( c_2 \). Similarly, the covariance acts linearly with respect to the second of its arguments.
14. Portfolio construction uses “weights” \( w_1, w_2, \ldots, w_N \) to denote the fractions of the portfolio invested in the different assets. Ordinarily “weights” are positive but in Math of Finance, negative values are used for for short positions and “weights” may be positive, zero or negative. The weight \( w_i \) is the amount invested in the \( i \)-th asset divided by the total value of the portfolio, and the sum of the weights obeys \( w_1 + w_2 + \ldots + w_N = 1 \). If the yearly return rates for the different assets are modeled as random variables \( r_1, \ldots, r_N \), then the yearly return rate for the portfolio is

\[
 r_P = w_1 r_1 + w_2 r_2 + \ldots + w_N r_N
\]

The weights \( w_1, \ldots, w_N \) do not depend on the random variables, and \( E[] \) acts linearly so the expected return rate for the portfolio is

\[
 \bar{r}_P \equiv E[r_P] = w_1 E[r_1] + w_2 E[r_2] + \ldots + w_N E[r_N] = w_1 \bar{r}_1 + w_2 \bar{r}_2 + \ldots + w_N \bar{r}_N
\]

The variance of the return rate for the portfolio is then

\[
 \sigma^2_P = \text{cov}(r_P, r_P) = \text{cov} \left( \sum_{i=1}^{N} w_i r_i, \sum_{j=1}^{N} w_j r_j \right) = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \text{cov}(r_i, r_j)
\]

15. Each choice of a set of weights \( \{w_1, w_N\} \) corresponds to a portfolio with volatility and expected return

\[
 \sigma_P(w_1, w_2, \ldots, w_N) = \sqrt{\sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \text{cov}(r_i, r_j)}
\]

\[
 \bar{r}_P(w_1, w_2, \ldots, w_N) = w_1 \bar{r}_1 + w_2 \bar{r}_2 + \ldots + w_N \bar{r}_N
\]

A volatility-expected return diagram is a plot of points \((\sigma_P, \bar{r}_P)\).

16. The weights \( w_i \) that minimize \( \sigma_P \) are the same as the weights that minimize \( \sigma^2_P \). The problem of minimizing \( \sigma_P \) is usually given as

\[
 \text{minimize } \sigma^2_P \quad \text{subject to } w_1 + w_2 + \ldots + w_N = 1
\]

in order to avoid working with the square root. The minimization problem can be solved by the method of Lagrange multipliers.