NYU Department of Politics - Quant I
Prof. Patrick Egan

Some Helpful Results Regarding the Math of Expectations

1 The basics: functions of the random variable Y

In all cases, Y is assumed to be a discrete random variable with a probability function $p(y)$ that accurately characterizes a population frequency distribution. With the exception of the final result, proofs are omitted.

1.1 Definitions

$$\begin{aligned} E(Y) &\equiv \sum_y y p(y) = \mu. \\ VAR(Y) &\equiv E[(Y-\mu)^2] = \sigma^2. \end{aligned}$$

1.2 Functions of a random variable

- Where $g(Y)$ is a real-valued function of Y,

$$E[g(Y)] = \sum_y g(y)p(y).$$

- Where $g_1, g_2 ... g_k$ are k real-valued functions of Y, we can "distribute expectations":

$$E[g_1(Y) + g_2(Y) + ... + g_k(Y)] = E[g_1(Y)] + E[g_2(Y)] + ... + E[g_k(Y)].$$

1.3 Constants

- Where c is a constant,

$$E(c) = c,$$

  and

$$E[cg(Y)] = cE[g(Y)].$$

- Population parameters, such as μ and σ^2, are constants.

1.4 The variance of a random variable

$$\sigma^2 = E(Y^2) - \mu^2.$$

Proof:

$$\begin{aligned}
\sigma^2 &= VAR(Y) \equiv E[(Y-\mu)^2] \\
&= E(Y^2 + \mu^2 - 2Y\mu) && \text{[Expanding the quadratic]} \\
&= E(Y^2) + E(\mu^2) - E(2Y\mu) && \text{[Distributing expectations]} \\
&= E(Y^2) + \mu^2 - 2\mu E(Y) && \text{[2, and } \mu \text{ (a population parameter) are constants]} \\
&= E(Y^2) + \mu^2 - 2\mu\mu && [E(Y) = \mu] \\
&= E(Y^2) - \mu^2 && \square.
\end{aligned}$$

Note this means we can also write

$$\sigma^2 = VAR(Y) = E(Y^2) - [E(Y)]^2.$$

2 More advanced: Functions of Random Variables $Y_1, Y_2 ... Y_N$

In all cases, $Y_1, Y_2 ... Y_N$ are assumed to be random variables with respective means μ_1, $\mu_2 ... \mu_N$ and variances σ_1^2, $\sigma_2^2 ... \sigma_N^2$.

2.1 Definitions

$$\begin{aligned}
COV(Y_1, Y_2) &\equiv E[(Y_1 - \mu_1)(Y_2 - \mu_2)]. \\
\text{correlation coefficient} &\equiv \rho_{Y_1 Y_2} \equiv \frac{COV(Y_1, Y_2)}{\sigma_1 \sigma_2}
\end{aligned}$$

2.2 Decomposing covariance

$$COV(Y_1, Y_2) = E(Y_1 Y_2) - E(Y_1)E(Y_2) = E(Y_1 Y_2) - \mu_1 \mu_2.$$

Proof:

$$\begin{aligned}
COV(Y_1, Y_2) &\equiv E[(Y_1 - \mu_1)(Y_2 - \mu_2)] \\
&= E[Y_1 Y_2 - Y_1 \mu_2 - Y_2 \mu_1 + \mu_1 \mu_2] \text{ (cross-multiplying)} \\
&= E(Y_1 Y_2) - E(Y_1 \mu_2) - E(Y_2 \mu_1) + E(\mu_1 \mu_2) \text{ (distributing expectations)} \\
&= E(Y_1 Y_2) - \mu_2 E(Y_1) - \mu_1 E(Y_2) + \mu_1 \mu_2 \text{ } (\mu_1, \mu_2 \text{ are constants}) \\
&= E(Y_1 Y_2) - \mu_2 \mu_1 - \mu_1 \mu_2 + \mu_1 \mu_2 \text{ (definition of } E(Y)) \\
&= E(Y_1 Y_2) - \mu_1 \mu_2. \ \square.
\end{aligned}$$

2.3 Covariance of independent random variables

- If Y_1, Y_2 independent, then

$$COV(Y_1, Y_2) = 0.$$

Proof:
From above,

$$COV(Y_1, Y_2) = E(Y_1 Y_2) - E(Y_1)E(Y_2).$$

But Y_1,Y_2 independent$\Rightarrow$ $E(Y_1 Y_2) = E(Y_1)E(Y_2)$, so

$$\begin{aligned} Y_1, Y_2 \text{ independent} &\Rightarrow COV(Y_1, Y_2) = E(Y_1)E(Y_2) - E(Y_1)E(Y_2) \\ &= 0. \end{aligned}$$

- However, the converse is not true. That is, $COV(Y_1, Y_2) = 0$ does not imply independence of Y_1, Y_2.

2.4 Expected value and variance of linear functions of random variables

- Consider U_1, a linear function of the random variables $Y_1, Y_2, ...Y_n$ and constants $a_1, a_2, ...a_n$,

$$U_1 = a_1 Y_1 + a_2 Y_2 + ... + a_n Y_n = \sum_{i=1}^{n} a_i Y_i,$$

and similarly

$$U_2 = \sum_{j=1}^{m} b_j X_j,$$

where $Y_1, Y_2, ...Y_n$ are random variables with $E(Y_i) = \mu_i$ and $X_1, X_2, ...X_n$ are random variables with $E(X_i) = \xi_i$ ["ksi-sub-i"]. Then (1), (2) and (3) below follow.

2.5 Expected value of a function of RVs

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

Proof:

$$\begin{aligned} E(U_1) &= E(a_1 Y_1) + E(a_2 Y_2) + ... + E(a_n Y_n) \text{ (distributing expections)} \\ &= a_1 E(Y_1) + a_2 E(Y_2) + ... + a_n E(Y_n) \text{ (factoring out constants)} \\ &= a_1 \mu_1 + a_2 \mu_2 + ... + a_n \mu_n \text{ (definition of expected value)} \\ &= \sum_{i=1}^{n} a_i \mu_i \quad \square. \end{aligned}$$

2.5.1 Variance of a function of RVs

$$VAR(U_1) = \sum_{i=1}^{n} a_i^2 VAR(Y_i) + 2\sum_{i<j} a_i a_j COV(Y_i, Y_j), \tag{2}$$

where the final sum is over all pairs (i, j) with $i < j$. (What does this mean in practice? That the covariance of each pair of RVs is taken only once under the summation sign.)

Proof:

$$\begin{aligned}
VAR(U_1) &\equiv E\left\{[U_1 - E(U_1)]^2\right\} \quad \text{[since } U_1 \text{ is itself a random variable]} \\
&= E\left[\left(\sum_{i=1}^{n} a_i Y_i - \sum_{i=1}^{n} a_i \mu_i\right)^2\right] \quad \text{(from above)} \\
&= E\left[\left(\sum_{i=1}^{n} a_i \left(Y_i - \mu_i\right)\right)^2\right] \quad \text{(factoring out constant)}
\end{aligned}$$

Note that the square of a sum always equals the sum of all the squares+ sum of all the (2 $\times$cross products). So for example

$$\begin{aligned}
(b_1 + b_2)^2 &= (b_1)^2 + (b_2)^2 + 2b_1 b_2 \\
(b_1 + b_2 + b_3)^2 &= (b_1)^2 + (b_2)^2 + (b_3)^2 + 2b_1 b_2 + 2b_1 b_3 + 2b_2 b_3 = \sum_{i=1}^{3} b_i^2 + 2\sum_{i<j}^{3} b_i b_j
\end{aligned}$$

and generally

$$\left(\sum_{i=1}^{n} b_i\right)^2 = (b_1 + b_2 + ... + b_n)^2 = \sum_{i=1}^{n} b_i^2 + 2\sum_{i<j}^{n} b_i b_j.$$

Now we can write

$$\begin{aligned}
E\left[\left(\sum_{i=1}^{n} a_i \left(Y_i - \mu_i\right)\right)^2\right] &= E\left[\sum_{i=1}^{n} a_i^2 \left(Y_i - \mu_i\right)^2 + \sum_{i<j}^{n} 2a_i a_j \left(Y_i - \mu_i\right)\left(Y_j - \mu_j\right)\right] \\
&= \sum_{i=1}^{n} a_i^2 E\left[\left(Y_i - \mu_i\right)^2\right] + \sum_{i<j} 2a_i a_j E\left[\left(Y_i - \mu_i\right)\left(Y_j - \mu_j\right)\right] \quad \text{(distributing expectations)} \\
&= \sum_{i=1}^{n} a_i^2 VAR(Y_i) + 2\sum_{i<j} a_i a_j COV(Y_i, Y_j). \quad \text{(def. of variance and covariance)} \ \square.
\end{aligned}$$

2.5.2 Covariance of two functions of RVs

$$COV(U_1, U_2) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j COV(Y_i, X_j). \tag{3}$$

Proof:

$$COV(U_1, U_2) = E\left[\left(\sum_{i=1}^{n} a_i Y_i - \sum_{i=1}^{n} a_i \mu_i\right)\left(\sum_{j=1}^{m} b_j X_j - \sum_{i=1}^{m} b_j \xi_i\right)\right]$$

(by def. of covariance, since U_1, U_2 are themselves RVs)

$$\begin{aligned}
&= E\left[\left(\sum_{i=1}^{n} a_i \left(Y_i - \mu_i\right)\right)\left(\sum_{j=1}^{m} b_j \left(X_j - \xi_i\right)\right)\right] \text{ (simplifying)} \\
&= E\left[\sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \left(Y_i - \mu_i\right)\left(X_j - \xi_i\right)\right] \text{ (cross-multiplying)} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j E\left[\left(Y_i - \mu_i\right)\left(X_j - \xi_i\right)\right] \text{ (distributing expectations)} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j COV(Y_i, X_j). \text{ (definition of covariance) } \square.
\end{aligned}$$

Ask yourself: observe that $COV(Y_i, Y_i) = VAR(Y_i)$. Do you see a link between statements (2) and (3) above?

2.6 The expected value and variance of the sample mean

- Now consider *independent* random variables $Y_1, Y_2, ...Y_n$ that have the *same* mean and the *same* variance, that is:

  $$E(Y_i) = \mu \text{ and } VAR(Y_i) = \sigma^2 \; \forall \; i.$$

  If we define the **sample mean** as the statistic

  $$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i,$$

  then

  $$E(\overline{Y}) = \mu \text{ and } VAR(\overline{Y}) = \frac{\sigma^2}{n}.$$

Proof:

Note that $\overline{Y}$ is a linear function of the independent random variables $Y_1, Y_2, ...Y_n$ with all constants a_i equal to $1/n$. So:

$$\begin{aligned}
E(\overline{Y}) &= E\left(\frac{1}{n}Y_1\right) + E\left(\frac{1}{n}Y_2\right) + ... + E\left(\frac{1}{n}Y_n\right) \\
&= \frac{1}{n}E(Y_1) + \frac{1}{n}E(Y_2) + ... + \frac{1}{n}E(Y_n) \\
&= \frac{1}{n}\mu + \frac{1}{n}\mu + ... + \frac{1}{n}\mu \\
&= \sum_{i=1}^{n} \frac{1}{n}\mu \\
&= \mu \quad \square.
\end{aligned}$$

And:

$$\begin{aligned}
VAR(\overline{Y}) &= \sum_{i=1}^{n} \left(\frac{1}{n}\right)^2 VAR(Y_i) + 2\sum_{i<j} \frac{1}{n}\frac{1}{n} COV(Y_i, Y_j) \text{ (from above)} \\
&= \sum_{i=1}^{n} \left(\frac{1}{n}\right)^2 VAR(Y_i) + 0 \quad \text{(independence } \Rightarrow COV(Y_i, Y_j) = 0 \forall Y_i, Y_j\text{)} \\
&= \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} \sigma^2 \text{ (factoring out constants, def. of variance)} \\
&= \left(\frac{1}{n}\right)^2 n\sigma^2 \\
&= \frac{\sigma^2}{n} \quad \square.
\end{aligned}$$