Expectations, Variances and Covariances of functions of one or more variables - week two

Luke Lloyd-Jones

Centre for Neurogenetics and Statistical Genomics, Queensland Brain Institute, University of Queensland, Brisbane, QLD, Australia

May 4, 2016

Overview

- Take a recap of last week and look at the excercises
- See how we can use the Taylor series trick to look at *variances* of complex variables
- Expectations and variances of products and ratios
- Derive the *sampling* variance for the regression coefficient

1 Exercises from last week

Exercises

1. Show (3) from properties of expectation and variances

   WTS

   $$\text{Var}(aX + b) = a^2\text{Var}(X)$$

   We know from the definition of variance that

   $$\text{Var}(X) = \mathbb{E}\{[X - \mathbb{E}(X)]^2\} = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

   Just plug in the $aX + b$ as X so

   $$\begin{aligned}
\text{Var}(aX + b) &= \mathbb{E}[(aX + b)^2] - [\mathbb{E}(aX + b)]^2 \\
&= \mathbb{E}[a^2X^2 + 2aXb + b^2] - [a\mathbb{E}(X) + b)]^2 \\
&= \mathbb{E}[a^2X^2 + 2abX + b^2] - a^2\mathbb{E}(X)^2 - 2ab\mathbb{E}(X) - b^2 \\
&= a^2\mathbb{E}(X^2) + 2ab\mathbb{E}(X) + b^2 - a^2\mathbb{E}(X)^2 - 2ab\mathbb{E}(X) - b^2 \\
&= a^2[\mathbb{E}(X^2) - \mathbb{E}(X)^2] \\
&= a^2\text{Var}(X)
\end{aligned}$$

2. Show (8) from properties of variances and covariances

   Want to show

   $$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$$

We will use the same trick of taking the definition and plugging in the wanted expression

$$
\begin{aligned}
\text{Var}(X+Y) &= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y) \\
&= \mathbb{E}[(X+Y)^2] - [\mathbb{E}(X+Y)]^2 \\
&= \mathbb{E}[X^2 + 2XY + Y^2] - [\mathbb{E}(X) + \mathbb{E}(Y)]^2 \\
&= \mathbb{E}[X^2] + 2\mathbb{E}[XY] + \mathbb{E}[Y^2] - \mathbb{E}(X)^2 - 2\mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(Y)^2
\end{aligned}
$$

Group terms and remember $\text{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$

$$
\begin{aligned}
&= \mathbb{E}[X^2] - \mathbb{E}(X)^2 + \mathbb{E}[Y^2] - \mathbb{E}(Y)^2 + 2[\mathbb{E}(XY) - 2\mathbb{E}(X)\mathbb{E}(Y)] \\
&= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y)
\end{aligned}
$$

3. Provide a *third* order approximation to the expectation of any random variable

$$
\begin{aligned}
\mathbb{E}[f(X)] &= \mathbb{E}\left[f(\mu_X) + (X-\mu_X)f'(\mu_X) + (X-\mu_X)^2\frac{f''(\mu_X)}{2} + (X-\mu_X)^3\frac{f'''(\mu_X)}{6} + \ldots\right] \\
&= \mathbb{E}[f(\mu_X)] + \mathbb{E}[(X-\mu_X)]f'(\mu_X) + \mathbb{E}[(X-\mu_X)^2]\frac{f''(\mu_X)}{2} + \mathbb{E}[(X-\mu_X)^3]\frac{f'''(\mu_X)}{6} + \ldots \\
&= \mathbb{E}[f(\mu_X)] + \sigma_X^2\frac{f''(\mu_X)}{2} + \mathbb{E}[(X-\mu_X)^3]\frac{f'''(\mu_X)}{6}
\end{aligned}
$$

4. Given your approximation in question (3) write down the expectation of $\log(X)$

$$\begin{aligned}\mathbb{E}[\log(X)] &= \log(\mu_X) + \sigma_X^2 \frac{f''(\mu_X)}{2} + \mathbb{E}[(X-\mu_X)^3]\frac{f'''(\mu_X)}{6} + \dots \\ &= \log(\mu_X) - \frac{\sigma_X^2}{2\mu_X^2} + \frac{\mathbb{E}[(X-\mu_X)^3]}{3\mu_X^3}\end{aligned}$$

5. If X is normally distributed what is the $\mathbb{E}(\log(X))$? If X is gamma distributed what is the $\mathbb{E}(\log(X))$? If X is normally distributed (doesn't really make sense) then $\mu_X = \mu$, $\sigma_X^2 = \sigma^2$, and $\mathbb{E}[(X-\mu_X)^3] = \mu^3 + 3\mu\sigma^2$. Plugging these in

$$\begin{aligned}\mathbb{E}[\log(X)] &= \log(\mu) - \frac{\sigma^2}{2\mu^2} + \frac{\mu^3 + 3\mu\sigma^2}{3\mu^3} \\ &= \log(\mu) - \frac{3\sigma^2}{6\mu^2} + \frac{2\mu^2 + 6\sigma^2}{6\mu^2} \\ &= \log(\mu) + \frac{2\mu^2 + 6\sigma^2 - 3\sigma^2}{6\mu^2} \\ &= \log(\mu) + \frac{2\mu^2 + 3\sigma^2}{6\mu^2}\end{aligned}$$

If X is gamma distributed then $\mu_X = \frac{\alpha}{\beta}$, $\sigma_X^2 = \frac{\alpha}{\beta^2}$. Plugging these in

$$\begin{aligned}\mathbb{E}[\log(X)] &= \log\left(\frac{\alpha}{\beta}\right) - \frac{\alpha}{\beta^2}\frac{1}{2\alpha^2/\beta^2} \\ &= \log\left(\frac{\alpha}{\beta}\right) - \frac{1}{2\alpha}\end{aligned}$$

You can try `digamma(100) - log(4)` and contrast it with the above approximation for $\alpha = 100$, $\beta = 4$.

6. Review the final derivation and contrast it with

https://www.youtube.com/watch?v=D1hgiAla3KI

2 Variances of complex variables

Last week we focused on expectations of complex variables and used this notion to look at classical unbiasedness of estimators. This week we will take a look at the same concept of using the Taylor series of a function to investigate *variances* of functions of random variables.

$$\text{Var}[f(X)] = \mathbb{E}\{[f(X) - \mathbb{E}[f(X)]]^2\}$$

We again just substitute for $f(X)$ the Taylor series approximation around μ_X

$$\begin{aligned}
\text{Var}[f(X)] = \mathbb{E}\Bigg\{\Bigg[\bigg(& f(\mu_X) + (X - \mu_X)f'(\mu_X) + (X - \mu_X)^2\frac{f''(\mu_X)}{2} + \dots\bigg) \\
& -\mathbb{E}\bigg(f(\mu_X) + (X - \mu_X)f'(\mu_X) + (X - \mu_X)^2\frac{f''(\mu_X)}{2} + \dots\bigg)\Bigg]^2\Bigg\} \\
& = \mathbb{E}\Bigg\{\Bigg[\bigg(f(\mu_X) + (X - \mu_X)f'(\mu_X) + (X - \mu_X)^2\frac{f''(\mu_X)}{2} + \dots\bigg) \\
& - \bigg(f(\mu_X) + \sigma_X^2\frac{f''(\mu_X)}{2} + \dots\bigg)\Bigg]^2\Bigg\}
\end{aligned}$$

Just keeping the first few terms

$$\approx \mathbb{E}\left\{\left[(X - \mu_X)f'(\mu_X) + [(X - \mu_X)^2 - \sigma_X^2]\frac{f''(\mu_X)}{2}\right]^2\right\}$$

Expand and take the expectation through the linear components

$$\approx \mathbb{E}\left\{(X-\mu_X)^2[f'(\mu_X)]^2 + 2(X-\mu_X)[f'(\mu_X)][(X-\mu_X)^2-\sigma_X^2]\frac{f''(\mu_X)}{2}\right.$$
$$\left. +[(X-\mu_X)^2-\sigma_X^2]^2\frac{f''(\mu_X)^2}{4}\right\}$$
$$\approx \sigma_X^2[f'(\mu_X)]^2 + 2\mathbb{E}[(X-\mu_X)^3]\frac{f'(\mu_X)f''(\mu_X)}{2} - 2\mathbb{E}[(X-\mu_X)]\sigma_X^2\frac{f'(\mu_X)f''(\mu_X)}{2}$$
$$+\{\mathbb{E}[(X-\mu_X)^4] - 2\mathbb{E}[(X-\mu_X)^2]\sigma_X^2 + \sigma_X^4\}\frac{f''(\mu_X)^2}{4}$$
$$\approx \sigma_X^2[f'(\mu_X)]^2 + 2\mathbb{E}[(X-\mu_X)^3]\frac{f'(\mu_X)f''(\mu_X)}{2} + \{\mathbb{E}[(X-\mu_X)^4] - \sigma_X^4\}\frac{f''(\mu_X)^2}{4}$$

Call $\mathbb{E}[(X-\mu_X)^3]$ and $\mathbb{E}[(X-\mu_X)^4]$ the third and fourth *moments* sometimes represented as μ_{3X} and μ_{4X}. There is an analogous representation for multiple random variables.

3 Expectations and variances of products and ratios

Let's consider the function $f(X) = h(X)g(X)$, where f is a function, which is a composite of two functions that may be random variables i.e., $h(X) = X$ or functions of random variables and not necessarily just one random variable.

Taking the first partial derivative of f with respect to h and g we have $\frac{\partial f(h,g)}{\partial h} = g$, $\frac{\partial f(h,g)}{\partial g} = h$, and $\frac{\partial^2 f(h,g)}{\partial h \partial g} = 1$, and all other partial derivatives are 0. Now we take these and plug it into our Taylor series approximation to the Expectation of two random variables equation (A1.4a) in **?** and we get

$$\mathbb{E}[f(h,g)] = f(\mu_h,\mu_g) + \sigma_h^2\frac{\partial^2 f}{2\partial h^2} + \sigma_g^2\frac{\partial^2 f}{2\partial g^2} + \sigma(h,g)\frac{\partial^2 f}{\partial h \partial g} \tag{1}$$

and we arrive at

$$\mathbb{E}[f(h,g)] = \mu_h \mu_g + \sigma(h,g) \tag{2}$$

If h and g were just random variables say X and Y then this expression would follow from the definition of covariance, where

$$\sigma(X,Y) = \mathbb{E}(XY) - \mu_X \mu_Y$$

We can now use the multivariable form of the variance equation, which is

$$= \mathbb{E}\left\{\left[(X-\mu_X)f'(\mu_X) + (Y-\mu_Y)f'(\mu_Y) + [(X-\mu_X)^2 - \sigma_X^2]\frac{f''(\mu_X)}{2}\right.\right. \tag{3}$$

$$\left.\left. + [(X-\mu_X)(Y-\mu_Y) - \sigma(X,Y)]\frac{\partial f^2(\mu_Y)}{\partial X \partial Y} + [(Y-\mu_Y)^2 - \sigma_Y^2]\frac{f''(\mu_Y)}{2}\right]^2\right\} \tag{4}$$

Substituting in our partial derivatives

$$\mathrm{Var}(XY) = \mathbb{E}\left\{[(X-\mu_X)\mu_Y + (Y-\mu_Y)\mu_X + (X-\mu_X)(Y-\mu_Y) - \sigma(X,Y)]^2\right\} \tag{5}$$

Expand and reduce (Exexcise 1)

The process is very similar for expectations and variances of ratios. Again if we treat $f(X,Y) = X/Y$ then the first and second order partial derivatives are

$$\frac{\partial f}{\partial X} = \frac{1}{Y} \tag{6}$$

$$\frac{\partial f}{\partial Y} = \frac{-X}{Y^2} \tag{7}$$

$$\frac{\partial^2 f}{\partial X^2} = 0 \tag{8}$$

$$\frac{\partial^2 f}{\partial Y^2} = \frac{2X}{Y^3} \tag{9}$$

$$\frac{\partial^2 f}{\partial Y \partial X} = \frac{-1}{Y^2} \tag{10}$$

So again with the expectation we have

$$\mathbb{E}[f(X,Y)] \approx f(\mu_X, \mu_Y) + \sigma_X^2 \frac{\partial^2 f}{2\partial X^2} + \sigma_Y^2 \frac{\partial^2 f}{2\partial Y^2} + \sigma(X,Y) \frac{\partial^2 f}{\partial X \partial Y} \tag{11}$$

$$\mathbb{E}[f(X,Y)] \approx \frac{\mu_X}{\mu_Y} \left[1 + \frac{2\sigma_Y^2}{\mu_Y^2} - \frac{\sigma(X,Y)}{\mu_X \mu_Y} \right] \tag{12}$$

And again for variances

$$\text{Var}[f(X,Y)] \approx \frac{\mu_X^2}{\mu_Y^2} \left[\frac{\sigma_X^2}{\mu_X^2} - \frac{2\sigma(X,Y)}{\mu_X \mu_Y} + \frac{\sigma_Y^2}{\mu_Y^2} \right] \tag{13}$$

These are both approximations

4 Derivation of sampling variances for regression and correlation coefficients

The least-squares regression coefficient is given by $\widehat{\beta} = \text{Cov}(U, V)/\text{Var}(U)$, where U is the independent variable. We want to know

(14)

$$\text{Var}(\widehat{\beta}) = \text{Var}\left(\frac{\text{Cov}(U, V)}{\text{Var}(U)}\right)$$

We would like to use the approximation

$$\text{Var}[f(X, Y)] \approx \frac{\mu_X^2}{\mu_Y^2}\left[\frac{\sigma_X^2}{\mu_X^2} - \frac{2\sigma(X, Y)}{\mu_X \mu_Y} + \frac{\sigma_Y^2}{\mu_Y^2}\right] \tag{15}$$

and we will require $\mu_X, \mu_Y, \sigma(X, Y), \sigma_X^2$, and σ_Y^2

$$\mu_X = \sigma(U, V) \tag{16}$$

$$\mu_Y = \text{Var}(U) \tag{17}$$

because the variance and covariance are unbiased estimators. We are going to make the further assumption that U and V are bivariate normal. Need

$$\sigma(X, Y) = \text{Cov}[\text{Cov}(U, V), \text{Var}(U)] \tag{18}$$

$$\sigma_X^2 = \text{Var}[\text{Cov}(U, V)] \tag{19}$$

$$\sigma_Y^2 = \text{Var}[\text{Var}(U)] \tag{20}$$

Assuming that U and V are bivariate normal we can use (A1.10b), (A1.14), and (A1.15) from **?**

$$\text{Var}[\text{Var}(X)] = \frac{2\sigma_X^2}{n} \tag{21}$$

$$\text{Var}[\text{Cov}(X,Y)] = \frac{\sigma_X^2\sigma_Y^2 + [\sigma(X,Y)]^2}{n} \tag{22}$$

(23)

$$\text{Cov}[\text{Cov}(X,Y), \text{Var}(X)] = \frac{2\sigma_X^2\sigma(X,Y)}{n}$$

and so

$$\sigma(X,Y) = \text{Cov}[\text{Cov}(U,V), \text{Var}(U)] = \frac{2\sigma_U^2\sigma(U,V)}{n} \tag{24}$$

$$\sigma_X^2 = \text{Var}[\text{Cov}(U,V)] = \frac{\sigma_U^2\sigma_V^2 + [\sigma(U,V)]^2}{n} \tag{25}$$

$$\sigma_Y^2 = \text{Var}[\text{Var}(U)] = \frac{2\sigma_U^4}{n} \tag{26}$$

Plugging these in

$$\text{Var}[\text{Cov}(U,V)/\text{Var}(U)] \approx \frac{[\sigma(U,V)]^2}{\sigma_U^4}\left[\frac{\text{Var}[\text{Cov}(U,V)]}{[\sigma(U,V)]^2} - \frac{2\frac{2\sigma_U^2\sigma(U,V)}{n}}{\sigma(U,V)\text{Var}(U)} + \frac{\frac{2\sigma_U^4}{n}}{[\text{Var}(U)]^2}\right] \tag{27}$$

$$\approx \frac{[\sigma(U,V)]^2}{\sigma_U^4}\left[\frac{\frac{\sigma_U^2\sigma_V^2+[\sigma(U,V)]^2}{n}}{[\sigma(U,V)]^2} - \frac{2\frac{2\sigma_U^2\sigma(U,V)}{n}}{\sigma(U,V)\sigma_U^2} + \frac{\frac{2\sigma_U^4}{n}}{[\sigma_U^2]^2}\right] \tag{28}$$

$$\approx \frac{[\sigma(U,V)]^2}{n\sigma_U^4}\left[\frac{\sigma_U^2\sigma_V^2+[\sigma(U,V)]^2}{[\sigma(U,V)]^2} - 4 + 2\right] \tag{29}$$

$$\approx \frac{1}{n\sigma_U^4}\left[\sigma_U^2\sigma_V^2+[\sigma(U,V)]^2 - 2[\sigma(U,V)]^2\right] \tag{30}$$

$$\approx \frac{\sigma_V^2}{n\sigma_U^2} - \frac{[\sigma(U,V)]^2}{n\sigma_U^4} \tag{31}$$

$$\approx \frac{\sigma_V^2}{n\sigma_U^2}\left[1 - \frac{[\sigma(U,V)]^2}{\sigma_V^2\sigma_U^2}\right] \tag{32}$$

$$\approx \frac{\sigma_V^2}{n\sigma_U^2}\left[1-\varrho^2\right] \tag{33}$$

Where ϱ is the correlation coefficient. This is what's in Lynch and Walsh. Compare this with the estimator for the standard error from first principles

$$\text{Var}(\widehat{\beta}) = \frac{\widehat{\sigma}^2}{ns_X}$$

Where s_X is the sample variance of X and $\widehat{\sigma}^2$ an estimate of the residual variance.

Exercises

1. Finish off expanding and reducing equation (**??**) and compare to equation (A1.18a)

2. Derive equation (**??**) using equation (A1.7c) in Lynch and Walsh

3. Review the final derivation and contrast it with that from first principles

4. Try and follow the same process of the final derivation to find (A1.20b) in Lynch and Walsh