Bivariate and Conditional Normals

The properties of the normal distribution are important for many LDV models. We have seen that if $Z \sim N(\mu, \Sigma)$, then the log of the joint density is

$$\ln f(Z) = -\frac{n}{2}\ln(2\pi) + \frac{1}{2}\ln(|\Sigma^{-1}|) - \frac{1}{2}(Z-\mu)'\Sigma^{-1}(Z-\mu)$$

Consider the special case of a bivariate standard-normal pair. That is,

$$Z = \begin{bmatrix} U \\ V \end{bmatrix} \mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

The bivariate ***independent*** standard normal pair is the restricted case where $\rho = 0$.

Bivariate and Conditional Normals

This covariance matrix has determinant $|\Sigma| = 1 - \rho^2$ and inverse

$$\Sigma^{-1} = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$$

Using these results, the log of the joint density for a bivariate standard normal pair is

$$\ln f(U,V) = -\ln(2\pi) - \frac{1}{2}\ln(1-\rho^2) - \frac{1}{2}\left(\frac{U^2 - 2\rho UV + V^2}{1-\rho^2}\right)$$

Bivariate and Conditional Normals

Since the marginals of a bivariate standard normal are univariate standard normal, we have

$$\ln f(V) = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}(V^2)$$

and the log of the conditonal density of U given V is

$$\begin{aligned} \ln f(U|V) &= \ln f(U,V) - \ln f(V) \\ &= -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(1-\rho^2) - \frac{1}{2}\frac{(U-\rho V)^2}{(1-\rho^2)} \end{aligned}$$

But this is just the log of a univariate normal density function. Hence, if (U,V) is bivariate standard normal, then $U|V \sim N[\rho V, (1-\rho^2)]$.

Interval Conditioning: Density Functions

- In this course, we will need the density function for a normal random variable conditional upon that variable falling in a restricted range.
- We will focus on upper tail conditioning (lower tail censoring). The reasoning is similar for lower tail conditioning.
- We will assume throught this section of material, that (U,V) is bivariate standard normal.
- The argument to the distribution or density function will be denoted v, the censoring threshold will be denoted K, and the random variables by V and U.

Interval Conditioning: Density Functions

1. In order to find the conditional distribution function for V given that it exceeds some cutoff K, we need the sets $A = \{V|V \leq v\}$, and $B = \{V|V > K\}$. Then,

$$\begin{aligned} F(v|V > K) &= P(V \leq v|V > K) = P(A|B) = P(AB)/P(B) \\ &= \begin{cases} \dfrac{\Phi(v) - \Phi(K)}{1 - \Phi(K)} & \text{for } v > K \\ 0 & \text{for } v \leq K \end{cases} \end{aligned}$$

This is because $AB = \{V|K < V \leq v\}$ for $v > K$, and $AB = \emptyset$ for $v \leq K$.

Interval Conditioning: Density Functions

The conditional density function is obtained through differentiation.

$$\begin{aligned} f(v|V > K) &= \frac{\partial F(v|V > K)}{\partial v} \\ &= \begin{cases} \dfrac{\phi(v)}{1 - \Phi(K)} & \text{for } v > K \\ 0 & \text{for } v < K \end{cases} \end{aligned}$$

Note that $F(v|V > K)$ is not differentiable at K.

Interval Conditioning: Density Functions

2. In order to find the conditional distribution function for U given that V exceeds some cutoff K, we need the sets $A = \{U|U \leq u\}$, and $B = \{V|V > K\}$. Then,

$$\begin{aligned} F(u|V > K) &= P(U \leq u|V > K) = P(A|B) = P(AB)/P(B) \\ &= \frac{\Phi(u, \infty) - \Phi(u, K)}{1 - \Phi(K)} \end{aligned}$$

Where we have used the fact that $P(A) = P(AB) + P(AB^c)$.

Interval Conditioning: Density Functions

The conditional density function is obtained through differentiation as

$$\begin{aligned} f(u|V > K) &= \frac{\partial F(u|V > K)}{\partial u} \\ &= \frac{\int_{-\infty}^{\infty} \phi(u, v) dv - \int_{-\infty}^{K} \phi(u, v) dv}{1 - \Phi(K)} \\ &= \frac{\int_{K}^{\infty} \phi(u, v) dv}{1 - \Phi(K)} \end{aligned}$$

Interval Conditioning: Expectations

- The results above may be used to determine the expected value of a random variable conditional on its falling in a certain interval.
- The results are valid for the case when (U, V) is bivariate standard normal.
- These conditional expectations form the basis for the two-stage estimations methods, and are also useful when signing the diagonal elements of the Hessian matrix.

Interval Conditioning: Expectations

1. The expected value of V given that it exceeds some cut-off K is

$$\begin{aligned} E(V|V>K) &= \int_K^\infty V f(V|V>K)dV \\ &= \int_K^\infty V \frac{\phi(V)}{1-\Phi(K)} dV \end{aligned}$$

Interval Conditioning: Expectations

The critical term in this integral is

$$\int_K^\infty V \exp(-\frac{1}{2}V^2)dV = -\exp(-\frac{1}{2}V^2)\Big|_K^\infty = \exp(-\frac{1}{2}K^2)$$

Hence

$$E(V|V>K) = \frac{\phi(K)}{1-\Phi(K)}$$

Interval Conditioning: Expectations

2. The expected value of U given that V exceeds some cut-off K is

$$\begin{aligned} E(U|V>K) &= \int_{-\infty}^\infty U f(U|V>K)dU \\ &= \int_{-\infty}^\infty U[\int_K^\infty \frac{\phi(U,V)}{1-\Phi(K)}dV]dU \\ &= \int_{-\infty}^\infty U[\int_K^\infty \frac{\phi(U|V)\phi(V)}{1-\Phi(K)}dV]dU \end{aligned}$$

and after interchanging the order of integration ...

Interval Conditioning: Expectations

$$
\begin{aligned}
E(U|V>K) &= \int_K^\infty E(U|V)\frac{\phi(V)}{1-\Phi(K)}dV \\
&= \rho E(V|V>K) \\
&= \rho\frac{\phi(K)}{1-\Phi(K)}
\end{aligned}
$$

where we have used the fact that $U|V \sim N(\rho V, 1-\rho^2)$ when (U,V) is bivariate standard normal.

Interval Conditioning: Expectations

3. The second moment of V given that it exceeds some cut-off K is

$$
\begin{aligned}
E(V^2|V>K) &= \int_K^\infty V^2 f(V|V>K)dV \\
&= \int_K^\infty V^2\frac{\phi(V)}{1-\Phi(K)}dV
\end{aligned}
$$

The critical term in this integral is

$$
\int_K^\infty V^2\exp(-\frac{1}{2}V^2)dV
$$

Interval Conditioning: Expectations

Solution will require integration by parts and L'Hôpital's rule. Letting $Z = V$ and $dW = V\exp(-\frac{1}{2}V^2)dV$, we have $dZ = dV$ and $W = -\exp(-\frac{1}{2}V^2)$. The integral is then

$$
\int_K^\infty V^2\exp(-\frac{1}{2}V^2)dV = -V\exp(-\frac{1}{2}V^2)\Big|_K^\infty + \int_K^\infty \exp(-\frac{1}{2}V^2)dV
$$

Interval Conditioning: Expectations

The limit of

$$
V\exp(-\frac{1}{2}V^2) = \frac{V}{\exp(\frac{1}{2}V^2)}
$$

is of the form $\frac{\infty}{\infty}$. Applying L'Hôpital's rule, the ratio of the derivatives is

$$
\frac{1}{V\exp(\frac{1}{2}V^2)}
$$

which has limit zero as V approaches ∞.

Interval Conditioning: Expectations

Hence

$$\int_K^\infty V^2 \exp(-\frac{1}{2}V^2)dV = K\exp(-\frac{1}{2}K^2) + \int_K^\infty \exp(-\frac{1}{2}V^2)dV$$

or

$$\int_K^\infty V^2\phi(V)dV = K\phi(K) + [1-\Phi(K)]$$

Using this result,

$$E(V^2|V>K) = 1 + K\left[\frac{\phi(K)}{1-\Phi(K)}\right]$$

Interval Conditioning: Expectations

The conditional variance is the second moment minus the square of the first, or

$$\begin{aligned}\sigma^2(V|V>K) &= 1 + K\left[\frac{\phi(K)}{1-\Phi(K)}\right] - \left[\frac{\phi(K)}{1-\Phi(K)}\right]^2 \\ &= 1 + \left[K - \frac{\phi(K)}{1-\Phi(K)}\right]\left[\frac{\phi(K)}{1-\Phi(K)}\right]\end{aligned}$$

Interval Conditioning: Expectations

4. The expected value of U^2 given that V exceeds some cut-off K is

$$\begin{aligned}E(U^2|V>K) &= \int_{-\infty}^\infty U^2 f(U|V>K)dU \\ &= \int_{-\infty}^\infty U^2[\int_K^\infty \frac{\phi(U,V)}{1-\Phi(K)}dV]dU \\ &= \int_{-\infty}^\infty U^2[\int_K^\infty \frac{\phi(U|V)\phi(V)}{1-\Phi(K)}dV]dU\end{aligned}$$

and after interchanging the order of integration ...

Interval Conditioning: Expectations

$$\begin{aligned}E(U^2|V>K) &= \int_K^\infty E(U^2|V)\frac{\phi(V)}{1-\Phi(K)}dV \\ &= \int_K^\infty [(1-\rho^2)+\rho^2V^2]\frac{\phi(V)}{1-\Phi(K)}dV\end{aligned}$$

where we have used the following:

- $U|V \sim N(\rho V, 1-\rho^2)$ when (U,V) is bivariate standard normal
- $E(U^2|V) = \sigma^2(U|V) + [E(U|V)]^2$

Interval Conditioning: Expectations

A little simplification gives

$$\begin{aligned} E(U^2|V > K) &= (1-\rho^2) + \rho^2 E(V^2|V > K) \\ &= (1-\rho^2) + \rho^2 \left[1 + K\frac{\phi(K)}{1-\Phi(K)}\right] \\ &= 1 + \rho^2 K \frac{\phi(K)}{1-\Phi(K)} \end{aligned}$$

Interval Conditioning: Expectations

The conditional variance is the second moment minus the square of the first, or

$$\begin{aligned} \sigma^2(U|V > K) &= 1 + \rho^2 K \frac{\phi(K)}{1-\Phi(K)} - \left[\frac{\rho\phi(K)}{1-\Phi(K)}\right]^2 \\ &= 1 + \rho^2 \left\{ \left[K - \frac{\phi(K)}{1-\Phi(K)}\right] \left[\frac{\phi(K)}{1-\Phi(K)}\right] \right\} \end{aligned}$$

Derivatives: Normal Distribution

In finding the score equations for many LDV models, we will need the partial derivatives of normal density and distribution functions. Let $\Phi(\cdot)$ and $\phi(\cdot)$ denote the standard normal distribution and density functions respectively. We have seen that

$$\ln \phi(Z) = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}Z^2$$

Thus,

1.
$$\frac{\partial \ln \phi(Z)}{\partial Z} = -Z$$

Derivatives: Normal Distribution

2.
$$\begin{aligned} \frac{\partial \phi(Z)}{\partial Z} &= \frac{\partial \ln \phi(Z)}{\partial Z}\phi(Z) \\ &= -Z\phi(Z) \end{aligned}$$

3.
$$\frac{\partial \Phi(Z)}{\partial Z} = \phi(Z)$$

by definition of a continuous density function and first fundamental theorem of integral calculus.

Derivatives: Normal Distribution

4.

$$\frac{\partial \ln \Phi(Z)}{\partial Z} = \frac{\partial \ln \Phi(Z)}{\partial \Phi(Z)} \frac{\partial \Phi(Z)}{\partial Z} = \frac{\phi(Z)}{\Phi(Z)}$$

5.

$$\frac{\partial \ln[1-\Phi(Z)]}{\partial Z} = \frac{\partial \ln[1-\Phi(Z)]}{\partial[1-\Phi(Z)]} \frac{\partial[1-\Phi(Z)]}{\partial Z} = -\frac{\phi(Z)}{1-\Phi(Z)}$$

Derivatives: Normal Distribution

6.

$$\frac{\partial \phi(Z)\Phi(Z)^{-1}}{\partial Z} = -Z\phi(Z)\Phi(Z)^{-1} - \phi(Z)^2\Phi(Z)^{-2} = -[Z + \frac{\phi(Z)}{\Phi(Z)}]\frac{\phi(Z)}{\Phi(Z)}$$

7.

$$\frac{\partial \phi(Z)[1-\Phi(Z)]^{-1}}{\partial Z} = -Z\phi(Z)[1-\Phi(Z)]^{-1} + \phi(Z)^2[1-\Phi(Z)]^{-2} = -[Z - \frac{\phi(Z)}{1-\Phi(Z)}]\frac{\phi(Z)}{1-\Phi(Z)}$$

Derivatives: Logistic Distribution

We will need similar partial derivatives for the Logit model. The Logistic distribution function is

$$\Psi(X) = \frac{e^X}{1+e^X} = e^X(1+e^X)^{-1}$$

The corresponding density function is

1.

$$\psi(X) = \frac{\partial \Psi(X)}{\partial X} = e^X(1+e^X)^{-1} - (e^X)^2(1+e^X)^{-2} = \Psi(X)[1-\Psi(X)]$$

Derivatives: Logistic Distribution

The second derivative of the Logistic distribution function is

2.

$$\frac{\partial^2 \Psi(X)}{\partial X^2} = \frac{\partial \psi(X)}{\partial X} = \frac{\partial \Psi(X)[1-\Psi(X)]}{\partial X} = \frac{\partial \Psi(X)}{\partial X}[1-\Psi(X)] + \Psi(X)\frac{\partial[1-\Psi(X)]}{\partial X} = \psi(X)[1-\Psi(X)] - \Psi(X)\psi(X) = \psi(X)[1-2\Psi(X)]$$

The logistic density function is similar to the standard normal density function, but has ”greater” tail areas. A logistically distributed random variable has zero mean and variance $\pi^2/3$.