Probability Review III

Harvard Math Camp - Econometrics

Ashesh Rambachan

Summer 2018

Outline

Outline

Useful Univariate Distributions

Not going to review them all in math camp but will refresh the most useful distributions. See the notes for a full review.

Bernoulli distribution

X is a discrete random variable that can only take on two values: $0, 1$. We write

$$f_X(x) = p^x(1-p)^{1-x}.$$

Note that

$$\begin{aligned} E[X^k] &= p, \quad k \geq 1 \\ V(X) &= p(1-p), \\ \mu_X(t) &= (1-p) + pe^t. \end{aligned}$$

X has a **Bernoulli distribution**.

Binomial distribution

X_i for $i = 1, \ldots, n$ are i.i.d Bernoulli random variables with $P(X_i = 1) = p$. Define

$$X = \sum_{i=1}^{n} X_i.$$

X follows a **binomial distribution** with parameters n and p. Takes values $1, 2, \ldots, n$ and

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

with

$$E[X] = np, \quad V(X) = np(1-p).$$

Uniform distribution

X is a continuous random variable with

$$f_X(x) = \frac{1}{b-a}$$

for $x \in [a, b]$ and 0 otherwise. X is **uniformly distributed on [a, b]** and write $X \sim U[a, b]$.

$$E[X] = \frac{1}{2}(a+b), \quad V(X) = \frac{1}{12}(b-a)^2.$$

Normal distribution

Suppose Z is continuously distributed with support over $\mathbb{R}$. X follows a **standard normal distribution** if

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

Denote it $Z \sim N(0,1)$ where $E[Z] = 0$, $V(Z) = 1$.

$X \sim N(\mu, \sigma^2)$ if

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

with $E[X] = \mu$, $V(X) = \sigma^2$ and $X = \mu + \sigma Z$, where $Z \sim N(0,1)$.

Normal distribution

The MGF of a standard normal random variable is incredibly useful. If $Z \sim N(0,1)$, then

$$M_Z(t) = e^{\frac{1}{2}t^2}.$$

If $X \sim N(\mu, \sigma^2)$, then

$$M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

Why?

Chi-squared Distribution

Let $Z_i \sim N(0,1)$ i.i.d. for $i = 1, \ldots, n$. Let

$$X = \sum_{i=1}^{n} Z_i^2.$$

X is a **chi-squared** random variable with n **degrees of freedom** and write $X \sim \chi_n^2$. Note

$$E[X] = n, \quad V(X) = 2n$$

.

Outline

The i.i.d. case

$Z = (Z_1, \ldots, Z_m)'$, where $Z_i \sim N(0,1)$ i.i.d. The joint density of Z is

$$
\begin{aligned}
f_Z(z) &= \Pi_{i=1}^m \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z_i^2} \\
&= (2\pi)^{n/2} \exp(-\frac{1}{2} z'z)
\end{aligned}
$$

Moreover, $E[Z] = 0$ and $V(Z) = I_m$.

The MGF of Z is

$$
\begin{aligned}
M_Z(t) &= E[e^{t'Z}] \\
&= \Pi_{i=1}^m E[e^{t_i z_i}] = e^{\frac{1}{2}t't}
\end{aligned}
$$

This is a useful reference point as we develop some results about the multivariate normal distribution.

Definition

The m-dimensional random vector X follows a m-**dimensional multivariate normal distribution** if and only if

$$a^T X$$

is normally distributed for all $a \in \mathbb{R}^m$.

We write $X \sim N_m(\mu, \Sigma)$, where $E[X] = \mu$ is the m-dimensional mean vector and $V(X) = \Sigma$ is the $m \times m$ dimensional covariance matrix.

What is its joint density? We use the following results to get there.

Density of Multivariate Normal

Result 1: Suppose $X \sim N(\mu, \Sigma)$. Then,

$$M_X(t) = e^{t'\mu + \frac{1}{2}t'\Sigma t}.$$

Proof: $t'X \sim N(t'\mu, t'\Sigma t)$. Therefore,

$$\begin{aligned} M_X(t) &= E[e^{t'X}] \\ &= E[e^Y], \quad Y \sim N(t'\mu, t'\Sigma t) \\ &= M_Y(1) \end{aligned}$$

Density of Multivariate Normal

Result 2: $X \sim N_m(\mu, \Sigma)$ and

$$Y = AX + b,$$

where $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$. Then,

$$Y \sim N_n(A\mu + b, A\Sigma A').$$

<u>Proof</u>: For $t \in \mathbb{R}^n$,

$$\begin{aligned}
M_Y(t) &= E[e^{t'Y}] \\
&= E[e^{t'(AX+b)}] \\
&= e^{t'b} E[e^{(A't)'X}] \\
&= e^{t'b} e^{(A't)'\mu + \frac{1}{2}(A't)'\Sigma(A't)'} \\
&= e^{t'(A\mu+b) + \frac{1}{2}t'(A\Sigma A')t}
\end{aligned}$$

Density of Multivariate Normal

We are now ready to derive the density of $X \sim N(\mu, \Sigma)$.

Suppose $X \sim N(\mu, \Sigma)$ and Σ has full column rank. Then, the density of X is given by

$$f_X(x) = (2\pi)^{-m/2}|\Sigma|^{-1/2}\exp(-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu))$$

Density of Multivariate Normal: Proof Sketch

Z is a m-dimensional random vector of i.i.d. standard normal random variables. We have

$$M_Z(t) = e^{\frac{1}{2}t't}$$

. so, $Z \sim N_m(0, I_m)$ with

$$f_Z(z) = (2\pi)^{-m/2} e^{-\frac{1}{2}z'z}$$

Let $X = \mu + \Sigma^{1/2} Z$. Using results, $X \sim N_m(\mu, \Sigma)$. From the multivariate transformation of random variables formula, we can get

$$f_X(x) = |\Sigma|^{-1/2} f_Z(\Sigma^{-1/2}(x - \mu))$$

Properties of Multivariate Normal Distribution

Next, we provide a list of a set of useful properties of the multivariate normal distribution. No need to memorize them but here so you're familiar with them.

- Results stated without proof.

Property #1: Concatenating independent multivariate normals

Property #1: If $X_1 \sim N_m(\mu_1, \Sigma_1), X_2 \sim N_n(\mu_2, \Sigma_2)$ and $X_1 \perp X_2$, then

$$X = (X_1', X_2')' \sim N_{m+n}(\mu, \Sigma)$$

where

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix}$$

Property #2: Subvectors are multivariate normals

Property #2: Let $X \sim N_m(\mu, \Sigma)$. Let X_1 be a p-dimensional sub-vector of X with $p < m$. Write

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

and

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Then, $X_1 \sim N_p(\mu_1, \Sigma_{11})$.

Property #3: $Cov(X_1, X_2) = 0 \iff X_1 \perp X_2$

Property #3: Let $X \sim N_m(\mu, \Sigma)$. Partition X into two sub-vectors. That is, write

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

and

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Then, $X_1 \perp X_2$ if and only if $\Sigma_{12} = \Sigma_{21} = 0$.

Property #4

Property #4: Let $X \sim N_m(\mu, \Sigma)$. If

$$Y = AX + b, \quad V = CX + d,$$

where $A, C \in \mathbb{R}^{n \times m}$ and $b, d \in \mathbb{R}^n$, then

$$Cov(Y, V) = A \Sigma C'.$$

Moreover, $Y \perp V$ if and only if

$$A \Sigma C' = 0.$$

Property #5: Linear conditional expectations

Property #5: Let $X \sim N_m(\mu, \Sigma)$ with $X = (X_1', X_2')'$, $\mu = (\mu_1', \mu_2')'$ and

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Provided that Σ_{22} has full rank, the conditional distribution of X_1 given $X_2 = x_2$ is

$$X_1 | X_2 = x_2 \sim N(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}).$$

Property #5: Linear Conditional Expectations

What's the intuition of this?

$$E[X_1|X_2 = x_2] = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2).$$

In 1-d, it becomes

$$E[X_1|X_2 = x_2] = E[X_1] + \frac{Cov(X_1, X_2)}{V(X_2)}(x_2 - E[X_2])$$

Next let's relabel $Y = X_1, X = X_2$ and re-arrange

$$E[Y|X = x] = (E[Y] - \frac{Cov(Y, X)}{V(X)}E[X]) + \frac{Cov(Y, X)}{V(X)}x.$$

This is simply the linear regression formula!

If (X, Y) are jointly normal, linear regression exactly returns the conditional expectation function.

Property #6: Quadratic Form of a Multivariate Normal

A **quadratic form** is a quantity of the form $y'Ay$, where A is a symmetric matrix.

Suppose that $Z_i \sim N(0,1)$ i.i.d. for $i = 1, \ldots, n$. We already know that $\sum_{i=1}^{n} Z_i^2 = Z'Z \sim \chi_n^2$.

Property #6:If $X \sim N_m(\mu, \Sigma)$ and Σ has full rank, then

$$(X - \mu)'\Sigma^{-1}(X - \mu) \sim \chi_m^2.$$

- Why? Let $Z = \Sigma^{-1/2}(X - \mu) \sim N_m(0, I_m)$. Then, $Z'Z \sim \chi_m^2$.