The Normal Distribution

Fall 2015
A. Stepanov
S.A. Culpepper

Univariate Normal Distribution

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

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

Important definitions and facts about the normal distribution:

1. The standard normal distribution is $Z \sim N(0,1)$.
2. The standard normal pdf is sometimes referred to as $f_Z(z) = \phi(z)$.
3. The standard normal cdf is, $\Phi(z) = F_Z(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}t^2} dt$.
4. Symmetry of the normal distribution implies $\Phi(z) = 1 - \Phi(-z)$.
5. $X \sim N(\mu, \sigma^2)$ has mgf $M_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$.

   Proof.

$$M_Z(t) = E(e^{Zt}) = \int_{-\infty}^{\infty} e^{zt} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2 - 2zt)} dz$$

$$= \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2}t^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2 - 2zt + t^2)} dz = e^{\frac{1}{2}t^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-t)^2} dz = e^{\frac{1}{2}t^2}$$

because the last integral was for a $Z \sim N(t, 1)$. Now make the transformation

$X = \sigma Z + \mu$,

$$M_X(t) = E(e^{Xt}) = E\left(e^{(\sigma Z + \mu)t}\right) = e^{\mu t} E(e^{Zt_*}) = e^{\mu t} E(e^{Zt_*}) = e^{\mu t} e^{\frac{1}{2}t_*^2}$$

$$= e^{\mu t + \frac{1}{2}\sigma^2 t^2}, t \in R$$

where $t_* = \sigma t$.

6. $Y = Z^2$ implies $Y \sim \chi^2(1)$.

Proof.

$$F_Y(y) = P(Y \le y) = P(Z^2 \le y) = P\left(-\sqrt{y} < Z \le \sqrt{y}\right)$$
$$= \Phi\left(\sqrt{y}\right) - \Phi\left(-\sqrt{y}\right) = 2\Phi\left(\sqrt{y}\right) - 1$$

$$f_Y(y) = F_Y'(y) = 2\phi\left(\sqrt{y}\right)\frac{1}{2\sqrt{y}} = \frac{y^{-\frac{1}{2}}}{\sqrt{2\pi}}e^{-\frac{1}{2}y}$$

7. If $X_i \sim N\left(\mu_i, \sigma_i^2\right)$ for $i = 1, \ldots, n$ are independent then for constants a_i $Y = \sum_{i=1}^n a_i X_i \sim N(\mu_Y, \sigma_Y^2)$ where $\mu_Y = \sum_{i=1}^n a_i \mu_i$ and $\sigma_Y^2 = \sum_{i=1}^n a_i^2 \sigma_i^2$.

Proof.

Using the independence assumption with mgfs implies,

$$M_Y(t) = E\left[e^{Yt}\right] = E\left[e^{\left(\sum_{i=1}^n a_i X_i\right)t}\right] = \prod_{i=1}^n E\left[e^{X_i(a_i t)}\right]$$
$$= \prod_{i=1}^n \exp\left[\mu_i a_i t + \frac{\sigma_i^2 a_i^2 t^2}{2}\right]$$
$$= \exp\left[\left(\sum_{i=1}^n a_i \mu_i\right)t + \frac{\left(\sum_{i=1}^n a_i^2 \sigma_i^2\right)t^2}{2}\right]$$

A. Stepanov
S.A. Culpepper

Bivariate Normal Distribution

$$f(x,y)=\frac{\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2-2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{x-\mu_x}{\sigma_y}\right)+\left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\}}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}},$$

$$-\infty<x<\infty,-\infty<y<\infty$$

8. We can say $\begin{bmatrix}X\\Y\end{bmatrix}\sim N_2(\boldsymbol{\mu},\boldsymbol{\Sigma})$,

$$\boldsymbol{\mu}=\begin{bmatrix}\mu_x\\\mu_y\end{bmatrix},\boldsymbol{\Sigma}=\begin{bmatrix}\sigma_x^2 & \rho\sigma_x\sigma_y\\\rho\sigma_x\sigma_y & \sigma_y^2\end{bmatrix}$$

9. $M_{X,Y}(t_x,t_y)=$
   $\exp\left[\mu_x t_x+\mu_y t_y+\frac{1}{2}\left(t_x^2\sigma_x^2+t_y^2\sigma_y^2+2\rho t_x t_y\sigma_x\sigma_y\right)\right],(t_x,t_y)\in R^2.$
10. The marginal distributions are $X\sim N(\mu_x,\sigma_x^2)$ and $Y\sim N(\mu_y,\sigma_y^2)$;
11. The correlation coefficient of X and Y is ρ; $\rho=0$ implies independence;
12. The distribution of $Y|X=x\sim N\left(\mu_y+\rho\frac{\sigma_y}{\sigma_x}(x-\mu_x),(1-\rho^2)\sigma_y^2\right)$.

Example

13. In a college health fitness program, let X denote the weight in kilograms of a male freshman at the beginning of the program and let Y denote his weight change during a semester. Assume that X and Y have a bivariate normal distribution with $\mu_x=75$, $\sigma_x=9$, $\mu_y=2.5$, $\sigma_y=1.5$, and $\rho=-0.6$. (lighter students tend to gain weight, while the heavier students tend to lose weight.) What proportion of the students that weigh 85 kg end up losing weight during the semester? That is, find $P(Y<0|X=85)$.

$$\mu_{y|x}=2.5-.6\frac{1.5}{9}(85-75)=1.5,\sigma_{y|x}^2=(1-.36)1.5^2=1.44,$$

$$Y|X=85\sim N(1.5,1.44)$$

$$P(Y<0|X=85)=P\left(Z<-\frac{1.5}{1.2}\right)=\Phi(-1.25)=0.1056$$

A. Stepanov
S.A. Culpepper

14. If X and Y are bivariate normal any linear combination is also normally distributed, $aX + bY \sim N\left(a\mu_x + b\mu_y, a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\rho\sigma_y\sigma_x\right)$ for all $(a, b) \in R$.

<u>Examples</u>

15. Reconsider the previous example. What proportion of the students weigh over 87 kg at the end of the semester? That is, find $P(X + Y > 87)$.

$$X + Y \sim N\left(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2 + 2\rho\sigma_y\sigma_x\right) \sim N(77.5, 67.05)$$

$$P(X + Y > 87) = P\left(Z > \frac{9.5}{8.1884}\right) = 1 - \Phi(1.16) = 0.1230$$

16. Suppose X and Y are standardized bivariate normal random variables with 0 means, unit variances, and correlation ρ. Show the distribution of $D = (X - Y)^2 \sim Gamma\left(\alpha = \frac{1}{2}, \theta = 4(1 - \rho)\right)$.

Proof. Note that $X - Y \sim N[0, 2(1 - \rho)]$. The mgf of D is,

$$M_D(t) = E\left[e^{Dt}\right] = E\left[e^{(X-Y)^2 t}\right] = E\left[e^{Z^2 t'}\right]$$

where $t' = 2(1 - \rho)t$. Recall $Z^2 \sim \chi_1^2$,

$$M_D(t) = \left(\frac{1}{1 - 2t'}\right)^{\frac{1}{2}} = \left(\frac{1}{1 - 4(1 - \rho)t}\right)^{\frac{1}{2}}, t < \frac{1}{4(1 - \rho)}$$

Multivariate Normal Distribution (MVN)

17. The MVN assumes all linear combinations of $X_1, \dots, X_n$ are normally distributed. The MVN pdf for $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is,

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})\right\}$$

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix}, \boldsymbol{\Sigma} = \mathbf{E}[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})'] = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{12} & \sigma_2^2 & & \sigma_{2n} \\ \vdots & & \ddots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \cdots & \sigma_n^2 \end{bmatrix}$$

where $\boldsymbol{\Sigma}$ is symmetric and non-negative definite.

Proof.

Let $\mathbf{t} = (t_1, \dots, t_n)'$ and $Y = \mathbf{t}'\mathbf{X} \sim N(\mu_y, \sigma_y^2)$ where $\mu_y = \mathbf{t}'\boldsymbol{\mu}$, $\sigma_y^2 = \mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}$, for all $\mathbf{t} \in R^n$. Recall that $M_y(t) = \exp\left(\mu_y t + \frac{1}{2}\sigma_y^2 t^2\right)$, which implies that $E(e^Y) = \exp\left(\mu_y + \frac{1}{2}\sigma_y^2\right)$. The multivariate mgf for $\mathbf{X}$ is,

$$M_{\mathbf{X}}(\mathbf{t}) = \mathrm{E}[\exp(\mathbf{t}'\mathbf{X})] = E(e^Y) = \exp\left(\mathbf{t}'\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}\right), \mathbf{t} \in R^n$$

Example.

18. $\mathbf{X} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \boldsymbol{\mu} = \begin{bmatrix} 5 \\ 3 \\ 7 \end{bmatrix}, \boldsymbol{\Sigma} = \begin{bmatrix} 4 & -1 & 0 \\ -1 & 4 & 2 \\ 0 & 2 & 9 \end{bmatrix}$

a. Find $P(X_1 > 6)$.

$X_1 \sim N(5, 4)$, $P(X_1 > 6) = P\left(Z > \frac{6-5}{\sqrt{4}}\right) = P(Z > 0.5) = \mathbf{0.3085}$.

b. Find $P(5X_2 + 4X_3 > 70)$.

$$E(5X_2 + 4X_3) = 43$$

$$Var(5X_2 + 4X_3) = [0 \quad 5 \quad 4]\begin{bmatrix} 4 & -1 & 0 \\ -1 & 4 & 2 \\ 0 & 2 & 9 \end{bmatrix}\begin{bmatrix} 0 \\ 5 \\ 4 \end{bmatrix} = 324.$$

$$P(5X_2 + 4X_3 > 70) = P(Z > 1.5) = 0.0668.$$

19. The standardized MVN for independent normal $Z_1, \ldots, Z_n$, $\mathbf{Z} \sim N_n(\mathbf{0}_n, \mathbf{I}_n)$, is,

$$f_{\mathbf{Z}}(\mathbf{z}) = \prod_{i=1}^{n} \phi(z_i) = \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2}\mathbf{z}'\mathbf{z}\right\}$$

Note the transformation $\mathbf{X} = \mathbf{\Sigma}^{\frac{1}{2}}\mathbf{Z} + \boldsymbol{\mu}$ also yields $f_{\mathbf{X}}(\mathbf{x})$.

20. For an m by n matrix $\mathbf{A}$ and m vector $\mathbf{b}$, $\mathbf{Y} = \mathbf{AX} + \mathbf{b} \sim N_m(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A\Sigma A}')$.

Proof.

$$M_{\mathbf{Y}}(\mathbf{t}) = E\{\exp[\mathbf{t}'\mathbf{Y}]\} = E\{\exp[\mathbf{t}'(\mathbf{AX} + \mathbf{b})]\} = \exp[\mathbf{t}'\mathbf{b}]\, E\{\exp[\mathbf{t}_*'\mathbf{X}]\}$$

where $\mathbf{t}_*' = \mathbf{t}'\mathbf{A}$, which yields,

$$M_{\mathbf{Y}}(\mathbf{t}) = \exp[\mathbf{t}'\mathbf{b}] \exp\left(\mathbf{t}_*'\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}_*'\mathbf{\Sigma}\mathbf{t}_*\right)$$

$$= \exp\left(\mathbf{t}'(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}) + \frac{1}{2}\mathbf{t}'\mathbf{A\Sigma A}'\mathbf{t}\right), \mathbf{t} \in R^m$$

21. The conditional distributions are also multivariate normal. We can write $\mathbf{X}$, $\boldsymbol{\mu}$, and $\mathbf{\Sigma}$ as,

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}, \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma_{11}} & \mathbf{\Sigma_{12}} \\ \mathbf{\Sigma_{12}'} & \mathbf{\Sigma_{22}} \end{bmatrix}$$

where $\mathbf{X}_1$ and $\boldsymbol{\mu}_1$ are $m \times 1$, $\mathbf{X}_2$ and $\boldsymbol{\mu}_2$ are $(n-m) \times 1$, $\mathbf{\Sigma_{11}}$ is $n \times n$, $\mathbf{\Sigma_{22}}$ is $(n-m) \times (n-m)$, and $\mathbf{\Sigma_{12}}$ is $m \times (n-m)$ defined as,

$$\mathbf{X}_1 = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}, \mathbf{X}_2 = \begin{bmatrix} x_{m+1} \\ \vdots \\ x_n \end{bmatrix}, \boldsymbol{\mu}_1 = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_m \end{bmatrix}, \mathbf{X}_2 = \begin{bmatrix} \mu_{m+1} \\ \vdots \\ \mu_n \end{bmatrix}$$

$$\mathbf{\Sigma}_{11} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{12} & \sigma_2^2 & & \sigma_{2m} \\ \vdots & & \ddots & \vdots \\ \sigma_{1m} & \sigma_{2m} & \cdots & \sigma_m^2 \end{bmatrix}, \mathbf{\Sigma}_{22} = \begin{bmatrix} \sigma_{m+1}^2 & \sigma_{m+1,m+2} & \cdots & \sigma_{m+1,n} \\ \sigma_{m+1,m+2} & \sigma_{m+2}^2 & & \sigma_{m+2,n} \\ \vdots & & \ddots & \vdots \\ \sigma_{m+1,n} & \sigma_{m+2,n} & \cdots & \sigma_n^2 \end{bmatrix}$$

$$\mathbf{\Sigma}_{12} = \begin{bmatrix} \sigma_{1,m+1} & \sigma_{1,m+2} & \cdots & \sigma_{1n} \\ \sigma_{2,m+1} & \sigma_{2,m+2} & & \sigma_{2n} \\ \vdots & & \ddots & \vdots \\ \sigma_{m,m+1} & \sigma_{m,m+2} & \cdots & \sigma_{m,n} \end{bmatrix}$$

It can be shown that $\mathbf{X_1}|\mathbf{X_2} \sim \boldsymbol{N}_m(\boldsymbol{\mu}_1 + \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2), \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{12}')$.

Example.

22. $\mathbf{X} \sim N_3(\boldsymbol{\mu}, \mathbf{\Sigma}), \boldsymbol{\mu} = \begin{bmatrix} 5 \\ 3 \\ 7 \end{bmatrix}, \mathbf{\Sigma} = \begin{bmatrix} 4 & -1 & 0 \\ -1 & 4 & 2 \\ 0 & 2 & 9 \end{bmatrix}$. Find $P(X_1 > 8|X_2 = 1, X_3 = 10)$.

$$\boldsymbol{\mu}_1 = 5, \boldsymbol{\mu}_2 = \begin{bmatrix} 3 \\ 7 \end{bmatrix}, \mathbf{\Sigma}_{11} = 4, \mathbf{\Sigma}_{22} = \begin{bmatrix} 4 & 2 \\ 2 & 9 \end{bmatrix}, \mathbf{\Sigma}_{22}^{-1} = \frac{1}{32}\begin{bmatrix} 9 & -2 \\ -2 & 4 \end{bmatrix}, \mathbf{\Sigma}_{12}' = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

$$E(X_1|X_2 = 1, X_3 = 10) = 5 + \frac{1}{32}\begin{bmatrix} -1 & 0 \end{bmatrix}\begin{bmatrix} 9 & -2 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} 1-3 \\ 10-7 \end{bmatrix} = 5.75$$

$$Var(X_1|X_2 = 1, X_3 = 10) = 4 + \frac{1}{32}\begin{bmatrix} -1 & 0 \end{bmatrix}\begin{bmatrix} 9 & -2 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} -1 \\ 0 \end{bmatrix} = 3.71875$$

We see that $X_1|X_2 = 1, X_3 = 10 \sim N(5.75, 3.71875)$, so $P(X_1 > 8|X_2 = 1, X_3 = 10) = P(Z > 1.17) = 0.1210$.