Math 442 – Mathematical Statistics **II**
Spring 2009
Solutions to the in-class assignments (2/1/09).

Definition: Two random variables X_1, X_2 are said to have Bivariate Normal distribution if their joint pdf is of the following form:

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi(1-\rho^2)^{1/2}\sigma_1\sigma_2} \cdot e^{-\frac{1}{2(1-\rho^2)}\cdot\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]} \text{ for } (x_1,x_2) \in \mathbb{R}^2.$$

The parameter space for the Bivariate Normal Distribution is $\mu_i \in \mathbb{R}$, $\sigma_i^2 > 0$ for $i = 1, 2$ and $\rho \in [-1, 1]$, and we use $(X_1, X_2) \sim BivariateNormal(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.

Theorem 1. *Let Z_1, Z_2 be independent and identically distributed $N(0,1)$. Define*

$$\begin{aligned} X_1 &:= \sigma_1 Z_1 + \mu_1 \quad \textit{and} \\ X_2 &:= \sigma_2[\rho Z_1 + (1-\rho^2)^{1/2} Z_2] + \mu_2. \end{aligned}$$

Then $(X_1, X_2) \sim BivariateNormal(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.

Proof. We note first that under this transformation,

$$Z_1 = \frac{X_1 - \mu_1}{\sigma_1} \quad \text{and} \quad Z_2 = \frac{\left(\frac{X_2-\mu_2}{\sigma_2}\right) - \rho\left(\frac{X_1-\mu_1}{\sigma_1}\right)}{(1-\rho^2)^{1/2}}.$$

The Jacobian of transform is, therefore,

$$\mathbf{J} = \begin{vmatrix} \frac{1}{\sigma_1} & 0 \\ \frac{-\rho}{\sigma_1(1-\rho^2)^{1/2}} & \frac{1}{\sigma_2(1-\rho^2)^{1/2}} \end{vmatrix} = \frac{1}{\sigma_1\sigma_2(1-\rho^2)^{1/2}}.$$

Given that $\sigma_1 > 0$ and $0 \le \rho^2 \le 1$, this is a positive constant. Thus we have, applying the transformation technique, that

$$\begin{aligned} f_{X_1,X_2}(x_1,x_2) &= \frac{1}{\sigma_1\sigma_2(1-\rho^2)^{1/2}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2} \cdot \sqrt{2\pi} e^{-\frac{1}{2}\frac{\left[\left(\frac{x_2-\mu_2}{\sigma_2}\right) - \rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\right]^2}{(1-\rho^2)}} \\ &= \frac{1}{2\pi(1-\rho^2)^{1/2}\sigma_1\sigma_2} \cdot e^{-\frac{1}{2(1-\rho^2)}\cdot\left[(1-\rho^2)\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \rho^2\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2\right]} \\ &= \frac{1}{2\pi(1-\rho^2)^{1/2}\sigma_1\sigma_2} \cdot e^{-\frac{1}{2(1-\rho^2)}\cdot\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]}. \end{aligned}$$

As the form of this joint distribution is precisely that of the Bivariate Normal distribution, it necessarily follows that $(X_1, X_2) \sim BivariateNormal(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ as desired. □

□

Theorem 2. *Let $(X_1, X_2) \sim BivariateNormal(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, then*

(a) $E(X_i) = \mu_i$, $Var(X_i) = \sigma_i^2$ and $Corr(X_1, X_2) = \rho$ for $i = 1, 2$,

(b) $X_i \sim N(\mu_i, \sigma_i^2)$ for $i = 1, 2$.

(c) $\rho = 0$ *if and only if* $X_1 \perp X_2$.

(d) $X_2\big|_{X_1=x_1} \sim N(\mu_2 + \rho\sigma_2(\frac{x_1-\mu_1}{\sigma_1}), (1-\rho^2)\sigma_2^2)$.

Proof. (a) We recall first a few properties of expected value, variance, and covariance.

- For $Y_1 \perp Y_2$:
$$E(aY_1 + bY_2 + c) = aE(Y_1) + bE(Y_2) + c$$
and
$$Var(aY_1 + bY_2 + c) = a^2Var(Y_2) + b^2Var(Y_2).$$

- For $Y_1, Y_2, \ldots, Y_m$ random variables such that $Cov(Y_i, Y_j)$ exists for all $i, j \in \{1, 2, \ldots, m\}$ and constants $a_1, \ldots, a_m, b_1, \ldots, b_m, c_1, c_2$:
$$Cov\big(c_1 + \sum_{i=1}^{m} a_iY_i, c_2 + \sum_{j=1}^{m} b_jY_j\big) = \sum_{i=1}^{m}\sum_{j=1}^{m} a_ib_jCov\big(Y_i, Y_j\big).$$

- For any random variable Y, $Cov(Y, Y) = E(YY) - E(Y)^2 = Var(Y)$, assuming said expected values exist.

Applying the above properties, noting $E(Z_i) = 0$ and $Var(Z_i) = 1$ for $i = 1, 2$, and as $Z_1 \perp Z_2$, $Cov(Z_1, Z_2) = 0$ we get:

$$\begin{aligned}
E(X_1) &= E(\sigma_1 Z_1 + \mu_1) \\
&= \sigma_1 E(Z_1) + \mu_1 = \mu_1, \\
E(X_2) &= E(\sigma_2[\rho Z_1 + (1-\rho^2)^{1/2}Z_2] + \mu_2) \\
&= \sigma_2\rho E(Z_1) + \sigma_2(1-\rho^2)^{1/2}E(Z_2) + \mu_2 = \mu_2, \\
Var(X_1) &= Var(\sigma_1 Z_1 + \mu_1) \\
&= \sigma_1^2 Var(Z_1) = \sigma_1^2, \\
Var(X_2) &= Var(\sigma_2[\rho Z_1 + (1-\rho^2)^{1/2}Z_2] + \mu_2) \\
&= \sigma_2^2\rho^2 Var(Z_1) + \sigma_2^2(1-\rho^2)Var(Z_2) = \sigma_2^2 \quad \text{and} \\
Cov(X_1, X_2) &= Cov(\sigma_1 Z_1 + \mu_1, \sigma_2\rho Z_1 + \sigma_2(1-\rho^2)^{1/2}Z_2 + \mu_2) \\
&= \sigma_1\sigma_2\rho Cov(Z_1, Z_1) + \sigma_1\sigma_2(1-\rho^2)^{1/2}Cov(Z_1, Z_2) \\
&= \sigma_1\sigma_2\rho Var(Z_1) = \sigma_1\sigma_2\rho.
\end{aligned}$$

It, then, follows that
$$Corr(X_1, X_2) = \frac{Cov(X_1, X_2)}{\sigma_1\sigma_2} = \frac{\sigma_1\sigma_2\rho}{\sigma_1\sigma_2} = \rho.$$

Thus we have established the required results for $E(X_i), Var(X_i)$, and $Corr(X_1, X_2)$ for $i = 1, 2$.

(b) We note that if $Y_1 \perp Y_2$, then $aY_1 \perp bY_2$ for constants a and b. We further recall some useful properties of moment generating functions.
Let $Y_1 \perp Y_2$, and a, b be constants. Then

$$M_{aY_1+b}(t) = e^{bt} \cdot M_{Y_1}(at) \quad \text{and}$$

$$M_{Y_1+Y_2}(t) = M_{Y_1}(t) \cdot M_{Y_2}(t).$$

As X_1 and X_2 are both linear combinations of Z_1 and Z_2 we use the above observations in order to find the marginal distributions of X_1 and X_2.

$$
\begin{aligned}
M_{X_1}(t) &= M_{\sigma_1 Z_1+\mu_1}(t) \\
&= e^{\mu_1 t}\cdot M_{Z_1}(\sigma_1 t) \\
&= e^{\mu_1 t}\cdot e^{\frac{1}{2}(\sigma_1 t)^2} \\
&= e^{\mu_1 t+\frac{1}{2}\sigma_1^2 t^2}.
\end{aligned}
$$

As this is the moment generating function of a normal distribution with mean μ_1 and variance σ_1^2, it follows that $X_1 \sim N(\mu_1, \sigma_1^2)$.

$$
\begin{aligned}
M_{X_2}(t) &= M_{\sigma_2[\rho Z_1+(1-\rho^2)^{1/2}Z_2]+\mu_2}(t) \\
&= e^{\mu_2 t}\cdot M_{\sigma_2\rho Z_1+\sigma_2(1-\rho^2)^{1/2}Z_2}(t) \\
&= e^{\mu_2 t}\cdot M_{Z_1}(\sigma_2\rho t) + M_{Z_2}(\sigma_2(1-\rho^2)^{1/2}t) \\
&= e^{\mu_2 t} e^{\frac{1}{2}(\sigma_2\rho t)^2} e^{\frac{1}{2}(\sigma_2(1-\rho^2)^{1/2}t)^2} \\
&= e^{\mu_2 t+\frac{1}{2}(\sigma_2^2\rho^2 t^2+\sigma_2^2(1-\rho^2)t^2)} \\
&= e^{\mu_2 t+\frac{1}{2}\sigma_2^2 t^2}
\end{aligned}
$$

This is the moment generating function of a normal distribution with mean μ_2 and variance σ_2^2. It now follows that $X_2 \sim N(\mu_2, \sigma_2^2)$. Thus we have shown $X_i \sim N(\mu_i, \sigma_i^2)$ for $i = 1, 2$.

(c) We know from Theorem 4.3.6 of the text, that if $X_1 \perp X_2$, then $Corr(X_1, X_2) = 0$ and as we have seen in Theorem 2.3, $Corr(X_1, X_2) = \rho$. It follow that if $X_1 \perp X_2$, then $\rho = 0$. Now suppose $\rho = 0$. Then the joint pdf of (X_1, X_2) has the following form:

$$
\begin{aligned}
f_{X_1,X_2}(x_1,x_2) &= \frac{1}{2\pi\sigma_1\sigma_2}\cdot e^{-\frac{1}{2}\cdot\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]} \quad \text{for } (x_1,x_2)\in\mathbb{R}^2 \\
&= \frac{1}{\sqrt{2\pi}\sigma_1}\cdot e^{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2}\cdot\frac{1}{\sqrt{2\pi}\sigma_2}\cdot e^{-\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2} \quad \text{for } x_1\in\mathbb{R} \quad \text{and} \quad x_2\in\mathbb{R} \\
&= f_{X_1}(X_1)\cdot f_{X_2}(X_2) \quad \text{for } x_1\in\mathbb{R} \quad \text{and} \quad x_2\in\mathbb{R}.
\end{aligned}
$$

As the joint pdf factors into a product of the two marginal pdfs (from Theorem 2.2), we see that $X_1 \perp X_2$. It follows that $\rho = 0$ if and only if $X_1 \perp X_2$.

(d) We again use properties of moment generate functions to obtain the conditional distribution of $X_2\big|_{X_1=x_1}$. Given $X_1 = x_1$, we have $Z_1 = \frac{x_1-\mu_1}{\sigma_1}$, and thus $X_2 = \sigma_2[\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right) + (1-\rho^2)^{1/2}Z_2] + \mu_2$. Let $\alpha = \sigma_2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right) + \mu_2$, and $\beta = \sigma_2(1-\rho^2)^{1/2}$, then $X_2\big|_{X_1=x_1} = \alpha + \beta Z_2$, and thus the moment generating function of X_2 given $X_1 = x_1$ is

$$
\begin{aligned}
M_{X_2\big|_{X_1=x_1}}(t) &= M_{\alpha+\beta Z_2}(t) \\
&= e^{\alpha t} M_{Z_2}(\beta t) \\
&= e^{\alpha t} e^{\frac{1}{2}\beta^2 t^2} \\
&= e^{\alpha t+\frac{1}{2}\beta^2 t^2}.
\end{aligned}
$$

This is the moment generating function of a normal distribution with mean α and variance β^2. Thus $X_2\big|_{X_1=x_1} \sim N(\alpha, \beta^2) = N(\mu_2 + \rho\sigma_2\left(\frac{x_1-\mu_1}{\sigma_1}\right), (1-\rho^2)\sigma_2^2)$ as desired.

□

Theorem 3. *Let* $(X_1, X_2) \sim BivariateNormal(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, *and define* $Y := aX_1 + bX_2 + c$. *Then* $Y \sim N(a\mu_1 + b\mu_2 + c, a^2\sigma_1^2 + b^2\sigma_2^2 + 2abCov(X_1, X_2))$.

Proof. As Y is a linear combination of X_1, X_2, which we have seen can be realized as linear combinations of Z_1 and Z_2, a pair of iid standard normal random variables, it follows that we can write Y itself as a linear combination of Z_1 and Z_2. That is,

$$\begin{aligned} X_1 &= \sigma_1 Z_1 + \mu_1, \\ X_2 &= \sigma_2[\rho Z_1 + (1-\rho^2)^{1/2} Z_2] + \mu_2 \quad \text{and} \\ Y &= aX_1 + bX_2 + c \\ &= a\big(\sigma_1 Z_1 + \mu_1\big) + b\big(\sigma_2[\rho Z_1 + (1-\rho^2)^{1/2} Z_2] + \mu_2\big) + c \\ &= (a\sigma_1 + b\sigma_2\rho)Z_1 + b\sigma_2(1-\rho^2)^{1/2} Z_2 + (a\mu_1 + b\mu_2 + c). \end{aligned}$$

As before, applying properties of moment generating functions we see that

$$\begin{aligned} M_Y(t) &= M_{(a\sigma_1+b\sigma_2\rho)Z_1+b\sigma_2(1-\rho^2)^{1/2}Z_2+(a\mu_1+b\mu_2+c)}(y) \\ &= e^{(a\mu_1+b\mu_2+c)t} M_{Z_1}((a\sigma_1 + b\sigma_2\rho)t) M_{Z_2}(b\sigma_2(1-\rho^2)^{1/2} t) \\ &= e^{(a\mu_1+b\mu_2+c)t} e^{\frac{1}{2}(a\sigma_1+b\sigma_2\rho)^2 t^2 + (b\sigma_2(1-\rho^2)^{1/2})^2 t^2} \\ &= e^{(a\mu_1+b\mu_2+c)t} e^{\frac{1}{2}\left(a^2\sigma_1^2+2ab\sigma_1\sigma_2\rho+b^2\sigma_2^2\rho^2+b^2\sigma_2^2(1-\rho^2)\right)t^2} \\ &= e^{(a\mu_1+b\mu_2+c)t+\frac{1}{2}\left(a^2\sigma_1^2+2ab\sigma_1\sigma_2\rho+b^2\sigma_2^2\right)t^2}. \end{aligned}$$

Recalling that $Cov(X_1, X_2) = \sigma_1\sigma_2\rho$ from our proof of Theorem 2.1, we see that

$$M_Y(t) = e^{(a\mu_1+b\mu_2+c)t+\frac{1}{2}\left(a^2\sigma_1^2+2abCov(X_1,X_2)+b^2\sigma_2^2\right)t^2}.$$

This is the moment generating function of a normal distribution with mean $a\mu_1 + b\mu_2 + c$ and variance $a^2\sigma_1^2 + 2abCov(X_1, X_2) + b^2\sigma_2^2$. It follows that $Y \sim N(a\mu_1 + b\mu_2 + c, a^2\sigma_1^2 + 2abCov(X_1, X_2) + b^2\sigma_2^2)$ as desired. □