**Review:** The derivative and integral of some important functions one should remember.

\[
\frac{d}{dx}(x^k) = kx^{k-1}
\]

\[
\frac{d}{dx}(e^x) = e^x
\]

\[
\frac{d}{dx} (\ln x) = \frac{1}{x}
\]

\[
\int_a^b x^k \, dx = \left( \frac{1}{k+1} x^{k+1} \right) \bigg|_a^b = \frac{1}{k+1} (b^{k+1} - a^{k+1})
\]

\[
\int_a^b e^x \, dx = e^x \bigg|_a^b = e^b - e^a
\]

**The Chain Rule**

\[
\frac{d}{dx} g[f(x)] = g'[f(x)] \cdot f'(x)
\]

For example: \( \frac{d}{dx} e^{x^2} = e^{x^2} \cdot 2x \)

**The Product Rule**

\[
\frac{d}{dx} [g(x) \cdot f(x)] = g'(x) f(x) + g(x) f'(x)
\]

**Review:** MGF, its second function: The m.g.f. will also generate the moments

Moment:

1st (population) moment:

\[
E(X) = \int x \cdot f(x) \, dx
\]

2nd (population) moment:

\[
E(X^2) = \int x^2 \cdot f(x) \, dx
\]

...
K\textsuperscript{th} (population) moment:

\[ E(X^k) = \int x^k \cdot f(x) \, dx \]

Take the K\textsuperscript{th} derivative of the \( M_X(t) \) with respect to \( t \), and the set \( t = 0 \), we obtain the K\textsuperscript{th} moment of \( X \) as follows:

\[
\left. \frac{d}{dt} M_X(t) \right|_{t=0} = E(X)
\]

\[
\left. \frac{d^2}{dt^2} M_X(t) \right|_{t=0} = E(X^2)
\]

... 

\[
\left. \frac{d^k}{dt^k} M_X(t) \right|_{t=0} = E(X^k)
\]

Note: The above general rules can be easily proven using calculus.

Example: When \( X \sim N(\mu, \sigma^2) \), we want to verify the above equations for \( k=1 \) & \( k=2 \).

\[
\frac{d}{dt} M_X(t) = (e^{\frac{\mu t}{2} + \frac{1}{2} \sigma^2 t^2}) \cdot (\mu + \sigma^2 t)
\]

(using the chain rule)

So when \( t=0 \)

\[
\left. \frac{d}{dt} M_X(t) \right|_{t=0} = \mu = E(X)
\]

\[
\frac{d^2}{dt^2} M_X(t)
\]

\[
= \frac{d}{dt} \left[ \frac{d}{dt} M_X(t) \right]
\]

\[
= \frac{d}{dt} \left[ (e^{\frac{\mu t}{2} + \frac{1}{2} \sigma^2 t^2}) \cdot (\mu + \sigma^2 t) \right]
\]

\[
= (e^{\frac{\mu t}{2} + \frac{1}{2} \sigma^2 t^2}) \cdot (\mu + \sigma^2 t)^2 + (e^{\frac{\mu t}{2} + \frac{1}{2} \sigma^2 t^2}) \cdot \sigma^2
\]

(using the result of the Product Rule)

And
\frac{d^2}{dt^2}M_X(t) \bigg|_{t=0} = \mu^2 + \sigma^2

= E(X^2)

Considering \( \sigma^2 = \text{Var}(X) = E(X^2) - \mu^2 \)

**Joint distribution, and independence**

**Definition.** The joint moment generating function of two random variables \( X \) and \( Y \) is defined as

\( M_{X,Y}(t_1, t_2) = E(e^{t_1X + t_2Y}) \)

**Theorem.** Two random variables \( X \) and \( Y \) are independent \( \iff \) (if and only if)

\( f_{X,Y}(x, y) = f_X(x)f_Y(y) \iff M_{X,Y}(t_1, t_2) = M_X(t_1)M_Y(t_2) \)

**Definition.** The covariance of two random variables \( X \) and \( Y \) is defined as

\( \text{COV}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \).

**Theorem.** If two random variables \( X \) and \( Y \) are independent, then we have \( \text{COV}(X, Y) = 0 \). (*Note: However, \( \text{COV}(X, Y) = 0 \) does not necessarily mean that \( X \) and \( Y \) are independent.)

**Exercise**

*Q1. Let \( X_1 \) & \( X_2 \) be independent \( N(\mu, \sigma^2) \). Prove that \( X_1 + X_2 \) and \( X_1 - X_2 \) are independent in two approaches: (1) pdf, (2) mgf*
Solution:

(1) The pdf approach: Let

\[ W = X_1 + X_2 \]
\[ Z = X_1 - X_2 \]

Then we have

\[ X_1 = \frac{W + Z}{2} \]
\[ X_2 = \frac{W - Z}{2} \]

This is a 1-1 transformation between \((X_1, X_2)\) and \((W, Z)\).

Define the Jacobian \( J \) of the transformation as:

\[
J = \begin{vmatrix}
\frac{\partial x_1}{\partial w} & \frac{\partial x_1}{\partial z} \\
\frac{\partial x_2}{\partial w} & \frac{\partial x_2}{\partial z}
\end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 2 & 2 \\ 1 & -1 \\ 2 & 2 \end{vmatrix} = -1/2
\]

Then the joint distribution of the new random variables is given by:

\[
f_{W,Z}(w, z) = f_{X_1,X_2}\left(\frac{w + z}{2}, \frac{w - z}{2}\right) |J|
\]

Since \(X_1\) and \(X_2\) are independent, we have:

\[
f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1)f_{X_2}(x_2)
\]

Thus we find:

\[
f_{W,Z}(w, z) = f_{X_1}\left(\frac{w + z}{2}\right)f_{X_2}\left(\frac{w - z}{2}\right)|J|
\]

\[
= \left(\frac{1}{\sqrt{2\pi}\sigma^2}\right)^2 e^{-\frac{1}{2\sigma^2}\left(\frac{w+z}{2} - \mu\right)^2 - \frac{1}{2\sigma^2}\left(\frac{w-z}{2} - \mu\right)^2} \cdot \frac{1}{2}
\]

\[
= \frac{1}{2} \left(\frac{1}{2\pi\sigma^2}\right) e^{-\frac{1}{4\sigma^2}(w-2\mu)^2 - \frac{1}{4\sigma^2}(z)^2}
\]

\[
= \frac{1}{\sqrt{2\pi} \times 2\sigma^2} e^{-\frac{1}{4\sigma^2}(w-2\mu)^2} \cdot \frac{1}{\sqrt{2\pi} \times 2\sigma^2} e^{-\frac{1}{4\sigma^2}z^2}
\]
So $Z$ and $W$ are independent. Furthermore we see that $X_1 + X_2 \sim N(2\mu, 2\sigma^2)$ and $X_1 - X_2 \sim N(0, 2\sigma^2)$

(2) The mgf approach: Now we prove that
\[ M_{X_1 + X_2, X_1 - X_2}(t_1, t_2) = M_{X_1 + X_2}(t_1)M_{X_1 - X_2}(t_2) \]

**Solution:**

Let $W = X_1 + X_2$ and $Z = X_1 - X_2$

* $W$ and $Z$ are independent if and only if $M_{W,Z}(t_1, t_2) = M_W(t_1)M_Z(t_2)$

\[ M_{W,Z}(t_1, t_2) = E(e^{t_1W + t_2Z}) \]

Substitute for $W$ and $Z$

\[ E(e^{t_1(X_1 + X_2) + t_2(X_1 - X_2)}) = E(e^{(t_1 + t_2)X_1 + (t_1 - t_2)X_2}) \]

\[ = E(e^{(t_1 + t_2)X_1})E(e^{(t_1 - t_2)X_2}) \]

because $X_1$ and $X_2$ are independent

* $E(e^{(t_1 + t_2)X_1}) = e^{\mu(t_1 + t_2) + \frac{1}{2}(t_1 + t_2)^2\sigma^2}$

\[ = e^{\mu t_1 + \mu t_2 + \frac{\sigma^2 t_1^2}{2} + \sigma^2 t_1 t_2 + \frac{\sigma^2 t_2^2}{2}} \]

* $E(e^{(t_1 - t_2)X_2}) = e^{\mu(t_1 - t_2) + \frac{1}{2}(t_1 - t_2)^2\sigma^2}$

\[ = e^{\mu t_1 - \mu t_2 + \frac{\sigma^2 t_1^2}{2} - \sigma^2 t_1 t_2 + \frac{\sigma^2 t_2^2}{2}} \]

\[ \therefore E(e^{(t_1 + t_2)X_1 + (t_1 - t_2)X_2}) = e^{\mu t_1 + \mu t_2 + \frac{\sigma^2 t_1^2}{2} + \sigma^2 t_1 t_2 + \frac{\sigma^2 t_2^2}{2} - \sigma^2 t_1 t_2 + \frac{\sigma^2 t_2^2}{2}} \]
\[
= \exp \left( \mu t_1 + \mu t_2 + \frac{\sigma^2 t_1^2}{2} + \sigma^2 t_1 t_2 + \frac{\sigma^2 t_2^2}{2} \right) \\
+ \mu t_1 - \mu t_2 + \frac{\sigma^2 t_1^2}{2} - \sigma^2 t_1 t_2 + \frac{\sigma^2 t_2^2}{2} \right)
\]

\[
= \exp \left( 2\mu t_1 + \frac{1}{2} \sigma^2 t_1^2 + \frac{1}{2} \sigma^2 t_2^2 \right)
\]

\[
= e^{2\mu t_1 + \frac{1}{2} \sigma^2 t_1^2} e^{\sigma^2 t_2^2 / 2}
\]

\[
= M_{X_1+X_2}(t_1) M_{X_1-X_2}(t_2)
\]

\[
= M_W(t_1) M_Z(t_2)
\]

\[
\therefore M_{W,Z}(t_1, t_2) = M_W(t_1) M_Z(t_2)
\]

\[
\therefore W = X_1 + X_2 \text{ and } Z = X_1 - X_2 \text{ are independent}
\]

**Q2.** Let \( Z \) be \( N(0,1) \), (1) please derive the covariance between \( Z \) and \( Z^2 \); (2) Are \( Z \) and \( Z^2 \) independent?
Solution:
(1) \( \text{cov}(Z, Z^2) = E \left[ (Z - E(Z)) \left( Z^2 - E(Z^2) \right) \right] \)
= \( E[(Z - 0)(Z^2 - 1)] = E[Z^3] = \frac{d^3}{dt^3}[\exp(\frac{1}{2}t^2)]|_{t=0} = 0 \)

(2) \( Z \) and \( Z^2 \) are not independent. You can do it in many ways, using (a) pdf, (b) cdf/probabilities, or (c) mgf.
For example, \( P(Z > 1, Z^2 > 1) = P(Z > 1) \neq P(Z > 1) \times P(Z^2 > 1) = 2[P(Z > 1)]^2 \)

Q3. Let \( X \sim N(3,4) \), please calculate \( P(1 < X < 3) \)

Solution:
\[
P(1 < X < 3) = P(\frac{1-3}{2} < \frac{X-3}{2} < \frac{3-3}{2})
= P(-1 < Z < 0)
= P(0 < Z < 1)
= 0.3413
\]
Note: In the above, \( Z \sim N(0,1) \)
Q4. Students A and B plans to meet at the SAC Seawolf market between 12noon and 1pm tomorrow. The one who arrives first will wait for the other for 30 minutes, and then leave. What is the chance that the two friends will be able to meet at SAC during their appointed time period assuming that each one will arrive at SAC independently, each at a random time between 12noon and 1pm?
**Solution:** $3/4$ as illustrated by the following figure.

Let $X$ and $Y$ denote the arrival time (in hour where 0 indicates 12noon and 1 indicates 1pm) of A and B respectively. Our assumption implies that $X$ and $Y$ each follows Uniform $[0,1]$ distribution, and furthermore, they are independent, Thus the joint pdf of $X$ and $Y$ follows the uniform distribution with $f(x,y) = 1$, if $(x,y)$ is in the square region, and $f(x,y) = 0$, outside the square region. A and B will be able to meet iff $|X - Y| \leq 1/2$ as represented by the red area $M$ bounded by the lines: $X - Y = -1/2$ and $X - Y = +1/2$.

![Diagram](image)

That is:

$$P(A \& B \text{ will meet}) = P\left(|X - Y| \leq \frac{1}{2}\right) = \int_{M} 1 \, dx \, dy = 1 - \frac{1}{8} - \frac{1}{8} = \frac{3}{4}$$

*Definitions: population correlation & sample correlation*

Definition: The population correlation coefficient $\rho$ is defined as:
\[ \rho = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X) \times \text{var}(Y)}} \]

Definition: Let \((X_1, Y_1), \ldots, (X_n, Y_n)\) be a random sample from a given bivariate population, then the sample correlation coefficient \(r\) is defined as:

\[ r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \]

*Definition: Bivariate Normal Random Variable*

\((X, Y) \sim BN(\mu_x, \sigma_x^2; \mu_y, \sigma_y^2; \rho)\) where \(\rho\) is the correlation between \(X \& Y\)

The joint p.d.f. of \((X, Y)\) is

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

Exercise: Please derive the mgf of the bivariate normal distribution.

**Q5.** Let \(X\) and \(Y\) be random variables with joint pdf

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

where \(-\infty < x < \infty, -\infty < y < \infty\). Then \(X\) and \(Y\) are said to have the bivariate normal distribution. The joint moment generating function for \(X\) and \(Y\) is

\[ M(t_1, t_2) = \exp \left[ t_1 \mu_x + t_2 \mu_y + \frac{1}{2} \left( t_1^2 \sigma_x^2 + 2\rho t_1 t_2 \sigma_x \sigma_y + t_2^2 \sigma_y^2 \right) \right]. \]

(a) Find the marginal pdf's of \(X\) and \(Y\);
(b) Prove that \( X \) and \( Y \) are independent if and only if \( \rho = 0 \).
(Here \( \rho \) is indeed, the <population\> correlation coefficient between \( X \) and \( Y \).
(c) Find the distribution of \((X + Y)\).
(d) Find the conditional pdf of \( f(x|y) \), and \( f(y|x) \)

**Solution:**

(a) The moment generating function of \( X \) can be given by

\[
M_X(t) = M(t, 0) = \exp \left[ \mu_X t + \frac{1}{2} \sigma_X^2 t^2 \right].
\]

Similarly, the moment generating function of \( Y \) can be given by

\[
M_Y(t) = M(0, t) = \exp \left[ \mu_Y t + \frac{1}{2} \sigma_Y^2 t^2 \right].
\]

Thus, \( X \) and \( Y \) are both marginally normal distributed, i.e.,

\( X \sim N(\mu_X, \sigma_X^2) \), and \( Y \sim N(\mu_Y, \sigma_Y^2) \).

The pdf of \( X \) is

\[
f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp \left[ -\frac{(x - \mu_X)^2}{2\sigma_X^2} \right].
\]

The pdf of \( Y \) is

\[
f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_Y} \exp \left[ -\frac{(y - \mu_Y)^2}{2\sigma_Y^2} \right].
\]

(b) If \( \rho = 0 \), then

\[
M(t_1, t_2) = \exp \left[ \mu_X t_1 + \mu_Y t_2 + \frac{1}{2} \left( \sigma_X^2 t_1^2 + \sigma_Y^2 t_2^2 \right) \right] = M(t_1, 0) \cdot M(0, t_2)
\]

Therefore, \( X \) and \( Y \) are independent.

If \( X \) and \( Y \) are independent, then
\[ M(t_1, t_2) = M(t_1, 0) \cdot M(0, t_2) = \exp \left[ \mu_x t_1 + \mu_y t_2 + \frac{1}{2} \left( \sigma_x^2 t_1^2 + \sigma_y^2 t_2^2 \right) \right] \]

\[ = \exp \left[ \mu_x t_1 + \mu_y t_2 + \frac{1}{2} \left( \sigma_x^2 t_1^2 + 2\rho \sigma_x \sigma_y t_1 t_2 + \sigma_y^2 t_2^2 \right) \right] \]

Therefore, \( \rho = 0 \)

(c) \[ M_{X+Y}(t) = E \left[ e^{t(X+Y)} \right] = E \left[ e^{tX+ty} \right] \]

Recall that \( M(t_1, t_2) = E \left[ e^{tX+x+y} \right] \), therefore we can obtain \( M_{X+Y}(t) \) by setting \( t_1 = t_2 = t \) in \( M(t_1, t_2) \).

That is,

\[ M_{X+Y}(t) = M(t, t) = \exp \left[ \mu_x t + \mu_y t + \frac{1}{2} \left( \sigma_x^2 t^2 + 2\rho \sigma_x \sigma_y t^2 + \sigma_y^2 t^2 \right) \right] \]

\[ = \exp \left[ (\mu_x + \mu_y) t + \frac{1}{2} \left( \sigma_x^2 + 2\rho \sigma_x \sigma_y + \sigma_y^2 \right) t^2 \right] \]

\[ \because X + Y \sim N(\mu = \mu_x + \mu_y, \sigma^2 = \sigma_x^2 + 2\rho \sigma_x \sigma_y + \sigma_y^2) \]

(d) The conditional distribution of \( X \) given \( Y = y \) is given by

\[ f(x | y) = \frac{f(x, y)}{f(y)} \]

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

Similarly, we have the conditional distribution of \( Y \) given \( X = x \) is
\[ f(y|x) = \frac{f(x,y)}{f(x)} \]
\[ = \frac{1}{\sqrt{2\pi \sigma_y \sqrt{1 - \rho^2}}} \exp \left\{ - \frac{\left( y - \mu_y - \frac{\sigma_y}{\sigma_x} \rho (x - \mu_x) \right)^2}{2(1 - \rho^2) \sigma_y^2} \right\}. \]

Therefore:

\[ X|Y = y \sim N \left( \mu_x + \rho \frac{\sigma_x}{\sigma_y} (y - \mu_y), (1 - \rho^2) \sigma_x^2 \right) \]
\[ Y|X = x \sim N \left( \mu_y + \rho \frac{\sigma_y}{\sigma_x} (x - \mu_x), (1 - \rho^2) \sigma_y^2 \right) \]

**Q6:** \( Z \sim N(0,1) \), and the value of random variable \( W \) depends on a coin (fair coin) flip: \( W = \begin{cases} Z, & \text{if head} \\ -Z, & \text{if tail} \end{cases} \)

find the distribution of \( W \). Is the joint distribution of \( Z \) and \( W \) a bivariate normal?
Answer:
\[ F_W(w) = P(W \leq w) \]
\[ = P(W \leq w | H)P(H) + P(W \leq w | T)P(T) \]
\[ = P(Z \leq w)P(H) + P(-Z \leq w)P(T) \]
\[ = F_Z(w) \times \frac{1}{2} + P(Z \geq -w) \times \frac{1}{2} \]
\[ = F_Z(w) \times \frac{1}{2} + P(Z \leq w) \times \frac{1}{2} \]
\[ = F_Z(w) + \frac{1}{2}F_Z(w) \times \frac{1}{2} \]
\[ = F_Z(w) \]

So \( W \sim N(0,1) \).

The joint distribution of \( Z \) and \( W \) is not normal. This can be shown by deriving the joint mgf of \( Z \) and \( W \) and compare it with the joint mgf of bivariate normal.

\[
M(t_1, t_2) = E(e^{t_1Z + t_2W})
\]
\[
= E(E(e^{t_1Z + t_2W} | \text{coin}))
\]
\[
= E(e^{t_1Z + t_2W} | H)P(H) + E(e^{t_1Z + t_2W} | T)P(T)
\]
\[
= E(e^{(t_1 + t_2)Z}) \times \frac{1}{2} + E(e^{(t_1 - t_2)Z}) \times \frac{1}{2}
\]
\[
= \frac{1}{2} \left( e^{\frac{(t_1 + t_2)^2}{2}} + e^{\frac{(t_1 - t_2)^2}{2}} \right)
\]

This is not the joint mgf of bivariate normal.

Alternatively, you can also derive the following probability:

\[
P(W + Z = 0)
\]
\[
= P(W + Z = 0 | H)P(H) + P(W + Z = 0 | T)P(T)
\]
\[
= P(2Z = 0) \times \frac{1}{2} + P(-Z + Z = 0) \times \frac{1}{2}
\]
\[
= 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{1}{2}
\]

This shows that \( W + Z \) cannot possibly follow a univariate normal distribution – which in turns shows that \( W \) and \( Z \) cannot possibly follow a bivariate normal distribution.