Economics 361
Problem Set #2 (Suggested Solutions)

Jun Ishii *
Department of Economics
Amherst College

Fall 2022

Question 1: Training Wheels

(a) Using the above, briefly explain why the *joint* distribution of X and Y is

$$f(x,y) \quad = \quad \frac{1}{2\Pi}\ e^{-\frac{1}{2}(x^2+y^2)}$$

ANS: X and Y are distributed *independently* of each other. Therefore, we know that $f(x,y) = f(x)\ \times\ f(y)$. Note that $\frac{1}{2\Pi}\ e^{-\frac{1}{2}(x^2+y^2)} = \frac{1}{\sqrt{2\Pi}}\ e^{-\frac{1}{2}x^2}\ \times\ \frac{1}{\sqrt{2\Pi}}\ e^{-\frac{1}{2}y^2}$.

(b) Fill in the ??? in the integral above.

ANS: The left most integral is for x. It indicates that x is unconstrained and can take any real value. This implies that the constraint $x + y \leq z$ must be applied to y. So y can take any value less than $z - x$. Therefore, ??? is $z - x$.

(c) Use the above to show that

$$f(z) \quad = \quad \frac{1}{2\Pi}\ e^{-\frac{1}{4}z^2} \int_{-\infty}^{\infty} e^{-(x-\frac{1}{2}z)^2}\ dx$$

ANS: This one is tricky. Let's start with the Fundamental Theorem of Calculus and work our way to Leibnitz's Rule

$$\begin{aligned} f(z) \quad &= \quad \frac{d}{dz}F_Z(z) \quad = \quad \frac{d}{dz}\int_{-\infty}^{\infty}\int_{-\infty}^{z-x} \underbrace{\frac{1}{2\Pi}\ e^{-\frac{1}{2}(x^2+y^2)}}_{f(x,y)}\ dy\ dx \\ &= \quad \frac{1}{2\Pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2}\ \underbrace{\frac{d}{dz}\left(\int_{-\infty}^{z-x} e^{-\frac{1}{2}y^2}\ dy\right)}_{(*)}\ dx \end{aligned}$$

*Office: Converse Hall 315 Phone: (413) 542-2901 E-mail: jishii@amherst.edu

This last step that interchanges the order of integration (with respect to x) and differentiation (with respect to z) is possible because z only appears in the domain of integration for y.

So we are now ready to apply Leibnitz's Rule to $(*)$. There is one tricky detail[1]: Leibnitz's Rule technically applies only for proper integrals. Consider the following:

$$(*) \quad = \quad \frac{d}{dz}\left(\int_{-\infty}^{z-x} e^{-\frac{1}{2}y^2}\, dy\right) \quad = \quad \underbrace{\frac{d}{dz}\left(\int_{-\infty}^{c} e^{-\frac{1}{2}y^2}\, dy\right)}_{\text{(A)}} + \underbrace{\frac{d}{dz}\left(\int_{c}^{z-x} e^{-\frac{1}{2}y^2}\, dy\right)}_{\text{(B)}}$$

We have decomposed $(*)$ into the derivative with respect to the improper integral (A) and the derivative with respect to the proper integral (B). c is some arbitrary finite constant less than $z-x$.

(A) is equal to zero as the improper integral in (A) converges to a finite constant that does not vary with z. We can apply Leibnitz's Rule to (B)

$$\begin{aligned} \frac{d}{dz}\left(\int_{c}^{z-x} e^{-\frac{1}{2}y^2}\, dy\right) \quad &= \quad e^{-\frac{1}{2}(z-x)^2}\underbrace{\frac{d(z-x)}{dz}}_{=1} - e^{-\frac{1}{2}c^2}\underbrace{\frac{d(c)}{dz}}_{=0} + \int_{c}^{z-x} \underbrace{\frac{d}{dz}\, e^{-\frac{1}{2}y^2}}_{=0}\, dy \\ &= \quad e^{-\frac{1}{2}(z-x)^2} \end{aligned}$$

Therefore

$$\begin{aligned} f(z) \quad = \quad \frac{1}{2\Pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2}\, e^{-\frac{1}{2}(z-x)^2}\, dx \quad &= \quad \frac{1}{2\Pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(2x^2+z^2-2zx)}\, dx \\ &= \quad \frac{1}{2\Pi}\int_{-\infty}^{\infty} e^{-(x^2+\frac{1}{2}z^2-zx)}\, dx \\ &= \quad \frac{1}{2\Pi}\, e^{-\frac{1}{4}z^2}\int_{-\infty}^{\infty} e^{-(x^2+\frac{1}{4}z^2-zx)}\, dx \\ &= \quad \frac{1}{2\Pi}\, e^{-\frac{1}{4}z^2}\int_{-\infty}^{\infty} e^{-(x-\frac{1}{2}z)^2}\, dx \end{aligned}$$

(d) Now, show that

$$\int_{-\infty}^{\infty} e^{-(x-\frac{1}{2}z)^2} \quad = \quad \sqrt{\Pi}$$

ANS: Note that the pdf of a Normal random variable X with mean μ and variance σ^2 is

$$f(x) \quad = \quad \frac{1}{\sqrt{2\Pi\sigma^2}}\, e^{-\frac{1}{2}\,\frac{(x-\mu)^2}{\sigma^2}}$$

So a Normal random variable X with mean $\frac{1}{2}z$ and variance $\frac{1}{2}$ has the pdf

$$f(x) \quad = \quad \frac{1}{\sqrt{2\;\Pi\;(\frac{1}{2})}}\, e^{-\frac{1}{2}\,\frac{(x-\frac{1}{2}z)^2}{\frac{1}{2}}} \quad = \quad \frac{1}{\sqrt{\Pi}}\, e^{-(x-\frac{1}{2}z)^2}$$

[1] A detail I tried to allow you to avoid in my later follow-up hint by allowing you to cheat and say $\frac{d}{dz}(\infty) = 0$

Note that for pdf $f(x)$

$$\int_{-\infty}^{\infty} f(x)\ dx \quad = \quad 1 \qquad \text{(This is essentially } P(S) = 1\text{)}$$

Therefore

$$\int_{-\infty}^{\infty} e^{-(x-\frac{1}{2}z)^2} \quad = \quad \sqrt{\Pi}\ \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{\Pi}}\ e^{-(x-\frac{1}{2}z)^2}\ dx}_{\int\ f(x)\ dx\ =\ 1} \quad = \quad \sqrt{\Pi}$$

(e) Use **(c)** and **(d)** to show that the pdf of Z is the pdf of a Normally distributed random variable. What are the mean and variance of Z?

ANS: Combining **(c)** and **(d)**

$$\begin{aligned} f(z) &= \frac{1}{2\Pi}\ e^{-\frac{1}{4}z^2} \int_{-\infty}^{\infty} e^{-(x-\frac{1}{2}z)^2}\ dx \\ &= \left(\frac{1}{2\Pi}\ e^{-\frac{1}{4}z^2}\right) \sqrt{\Pi} \\ &= \frac{1}{\sqrt{2\Pi\ \cdot\ 2}}\ e^{-\frac{1}{2}\ \frac{(z-0)^2}{2}} \end{aligned}$$

Note that this is the pdf for a Normal random variable with mean zero and variance 2, N(0,2).

General Note: The steps **(a)** - **(e)** above can be used to show the following more general result:

> Let X be distributed $N(\mu_X, \sigma_X^2)$ and Y distributed $N(\mu_Y, \sigma_Y^2)$. If X and Y are distributed independently of each other, then Z is distributed $N(\mu_x + \mu_Y, \sigma_X^2 + \sigma_Y^2)$

Question 2: Solo Time

Let X be a random variable distributed Standard Normal, N(0,1). Show that the pdf of $Z \equiv X^2$ is the same as that of a random variable distributed chi-squared with one degree of freedom (χ_1^2).

ANS: Let us start with the hint

$$\begin{aligned} F_Z(z) \quad = \quad P_Z(\{Z \le z\}) \quad &= \quad P_X(\{-\sqrt{z} \le X \le \sqrt{z}\}) \\ &\quad\ \text{for } z \ge 0 \\ &= \quad \int_{-\sqrt{z}}^{\sqrt{z}} \underbrace{\frac{1}{\sqrt{2\Pi}}\, e^{-\frac{1}{2}\, x^2}}_{f(x)}\, dx \end{aligned}$$

Using the Fundamental Theorem of Calculus and Leibnitz's Rule

$$\begin{aligned} f(z) \quad &= \quad \frac{d}{dz} F_Z(z) \quad = \quad \frac{d}{dz} \int_{-\sqrt{z}}^{\sqrt{z}} \underbrace{\frac{1}{\sqrt{2\Pi}}\, e^{-\frac{1}{2}\, x^2}}_{f(x)}\, dx \\ &= \quad \frac{1}{\sqrt{2\Pi}}\, e^{-\frac{1}{2}(\sqrt{z})^2}\, \underbrace{\frac{d}{dz}(\sqrt{z})}_{=\frac{1}{2}\,\frac{1}{\sqrt{z}}} \; - \; \frac{1}{\sqrt{2\Pi}}\, e^{-\frac{1}{2}(-\sqrt{z})^2}\, \underbrace{\frac{d}{dz}(-\sqrt{z})}_{=-\frac{1}{2}\,\frac{1}{\sqrt{z}}} \; + \; \int_{-\sqrt{z}}^{\sqrt{z}} \underbrace{\frac{d}{dz}\left(\frac{1}{\sqrt{2\Pi}}\, e^{-\frac{1}{2}x^2}\right)}_{=0} dx \\ &= \quad \frac{1}{\sqrt{2\Pi}}\, e^{-\frac{1}{2}z}\, \frac{1}{\sqrt{z}} \qquad \text{for } z \ge 0 \end{aligned}$$

Question 3: Simple Change of Variables

Let $Y \equiv a + b\ X$ where a, b are known real valued constants.

Use the change of variables technique to show that the pdf of Y is the same as that of a random variable distributed Normal with mean $a + b\mu$ and variance $b^2\sigma^2$, $\mathrm{N}(a + b\mu, b^2\sigma^2)$

ANS: This question is a straightforward application of the above change of variables technique. Note that $Y = g(X) = a + bX$. So the inverse function is $g^{-1}(Y) = \frac{Y-a}{b}$

$$\begin{aligned}
f(y) &= f_X(g^{-1}(y)) \left|\frac{d}{dy}\ g^{-1}(y)\right| \\
&= f_X(\frac{y-a}{b}) \left|\frac{d}{dy}\ \frac{y-a}{b}\right| \\
&= \frac{1}{\sqrt{2\Pi\sigma^2}}\ e^{-\frac{1}{2}\ \frac{(\frac{y-a}{b}-\mu)^2}{\sigma^2}}\ \left|\frac{1}{b}\right| \\
&= \frac{1}{\sqrt{2\Pi b^2\sigma^2}}\ e^{-\frac{1}{2}\frac{(y-(a+b\mu))^2}{b^2\sigma^2}}
\end{aligned}$$

The last step shows that the pdf of Y is the same as a random variable distributed $\mathrm{N}(a+b\mu, b^2\sigma^2)$.

Note: $\left|\frac{1}{b}\right| = \frac{1}{\sqrt{b^2}}$

BONUS: Use change of variables to show the result in Question 2: the square of a random variable distributed N(0,1) is itself distributed χ^2_1.

ANS: Recall the standard normal, $Z \sim N(0,1)$. Change of variables can be used to show that $Y \equiv Z^2$ is distributed Chi-squared with 1 degree of freedom, $Z \sim \chi^2_1$

Let us start with the Gaussian integral associated with Z

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\Pi}} e^{-\frac{1}{2}z^2} dz = \int_{-\infty}^{0} \frac{1}{\sqrt{2\Pi}} e^{-\frac{1}{2}z^2} dz + \int_{0}^{+\infty} \frac{1}{\sqrt{2\Pi}} e^{-\frac{1}{2}z^2} dz$$

Note that Y is a monotonic transformation of Z only for non-negative or non-positive domains of Z. So we split the Gaussian integral into the two appropriate halves.

$$\begin{aligned}
\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\Pi}}\ e^{-\frac{1}{2}z^2} dz &= \int_{-\infty}^{0} \frac{1}{\sqrt{2\Pi}}\ e^{-\frac{1}{2}(\sqrt{y})^2}\ \left|\frac{dz(y)}{dy}\right|\ dy\ +\ \int_{0}^{+\infty} \frac{1}{\sqrt{2\Pi}}\ e^{-\frac{1}{2}(\sqrt{y})^2}\ \left|\frac{dz(y)}{dy}\right|\ dy \\
&= 2\ \times\ \int_{0}^{+\infty} \frac{1}{\sqrt{2\Pi}}\ e^{-\frac{1}{2}(\sqrt{y})^2}\ \frac{1}{2\sqrt{y}}\ dy\ \ \text{(by symmetry)} \\
&= \int_{0}^{+\infty} \frac{1}{\sqrt{2\Pi}}\ e^{-\frac{1}{2}y}\ \frac{1}{\sqrt{y}}\ dy
\end{aligned}$$

The pdf of a χ^2_1 random variable is

$$f(y) = \frac{1}{\Gamma(1/2)2^{1/2}}\ y^{1/2-1}\ e^{-y/2} \quad \text{for}\ \ y \in (0, +\infty)$$

We can show that $\Gamma(1/2) = \sqrt{\Pi}$.[2] So

$$f(y) = \frac{1}{\sqrt{2\Pi}}\ e^{-\frac{1}{2}y}\ \frac{1}{\sqrt{y}} \quad \text{for}\ \ y \in (0, +\infty)$$

which is the integrand from the above change of variables from Z to Y.

NOTE: If you do not understand the above answer to the BONUS question, do not worry about it

[2] Apply change of variables, $u = 2\sqrt{y}$, on the Gamma function and use the Gaussian integral result: $\Gamma(1/2) = \sqrt{2}\ \int_0^{+\infty} e^{-\frac{1}{2}u^2} du = \sqrt{2}\ \frac{\sqrt{2\Pi}}{2} = \sqrt{\Pi}$

Question 4: Not Quite Normal

The pdf of the log Normal (for given parameter values α and β) is

$$f_X(x) = \frac{1}{\sqrt{2\pi\beta}} \frac{e^{\frac{-(ln(x)-\alpha)^2}{2\beta}}}{x} \quad \text{for } x \geq 0$$

α may take any real value but β must be positive.

Consider a random variable Y that is distributed Normal with mean μ and variance σ^2, $N(\mu, \sigma^2)$.

(a) Show that the pdf of random variable $X \equiv e^Y$ is distributed log Normal. State what (α, β) are in terms of (μ, σ^2)

ANS: Just use the "Change of Variables" technique introduced in Q3. Note that $X \equiv e^Y$ is a monotonic transformation of Y.

$$\begin{aligned} f_X(x) = f(x) &= f_Y(g^{-1}(x)) \left| \frac{d}{dx} g^{-1}(x) \right| \\ &\text{where} \\ f_Y(y) &= \frac{1}{\sqrt{2\Pi\sigma^2}} e^{-\frac{(y-\mu))^2}{2\sigma^2}} \\ g^{-1}(x) &= ln(x) \qquad \text{and thus} \left| \frac{d}{dx} g^{-1}(x) \right| = \frac{1}{x} \quad \text{for } x > 0 \end{aligned}$$

You should get the log Normal distribution with $\alpha = \mu$ and $\beta = \sigma^2$

(b) Briefly explain why the log Normal distribution is called, well, the log Normal distribution!

ANS: Think in terms of $g^{-1}(x)$. The natural log transformation of a random variable distributed log Normal results in a random variable that is distributed Normal! Thus, the log Normal random variable is a variable that is "distributed Normal in the (natural) logarithm."

(c) Based on what you know about the shape of the Normal pdf (when graphing $f(y)$ against y), provide an intuitive reason why the log Normal pdf is uni-modal but not symmetric.

ANS: There are several possible answers. Here is one:

Consider the graph of the pdf for the Normal distribution (either from lecture notes or assigned readings), which unimodal and symmetric around the mean (μ).

- The density (height of pdf) of the Normal random variable rises from $y \to -\infty$ to $y = \mu$ and then falls from $y = \mu$ to $y \to -\infty$. This indicates that the likelihood of associated realizations are increasing from $(-\infty, \mu)$ and decreasing from $(\mu, +\infty)$. But note that $X \equiv e^Y$ is not only a monotonic transformation of Y but also a one-to-one transformation. For each value x there is exactly one corresponding value of y and vice versa. Therefore, the likelihood must be increasing from $x = e^{-\infty} \to 0$ to $x = e^{\mu}$ and then decreasing thereafter. So, the log Normal distribution must be unimodal around e^{μ}.

- But note that, for the Normal random variable Y, half the probability mass lies between $(-\infty, \mu)$ and half between $(\mu, +\infty)$. As probability masses must match-up, half the probability mass for log Normal random variable X must lie between $(0, e^{\mu})$ and half between $(e^{\mu}, +\infty)$. But those two ranges are not equidistant, suggesting that the log Normal distribution is not symmetric.

(d) Based on our discussion so far, what might a stock analyst be predicting about stock price when s/he claims that s/he is "betting on a far tail outcome."

ANS: The tail that goes the farthest is the right tail, representing large realizations of X. So the stock analyst is betting that the stock price will be high.

(e) Consider a second random variable, Z, that is also distributed log Normal. Further, let X and Z be distributed independently of each other. Recall the result we derived in Question 1. Intuitively, explain why $W \equiv X \times Z$ is also distributed log Normal. (You do not need to show formally, although you could using the tools we have developed in this problem set)

ANS: It is actually not that hard to derive the result formally – well within your grasp if you were able to do the earlier questions. There are several possible answers (as there are several intuitive paths). Here is one:

Let $F = ln(X)$ and $G = ln(Z)$. From above, we know that F and G must each be distributed Normally. Additionally, as X and Z are distributed independently of each other, F and G must be too – intuitively, if X and Y are not (statistically) informative of each other, how can their log transformations be? We showed earlier that the sum of independent Normal random variables is normal too. Let $H = F + G$. Consider e^H. From above, we know that e^H must be distributed log Normal. But $e^H = e^{F+G} = e^F\ e^G = X \times Z = W$ So W must be distributed log Normal.

Question 5: Patent Race, Part I

Consider two research and development (R&D) firms who are both working on an improved pain reliever. Let us dub the two firms "incumbent" (**I**) and "entrant" (**E**).

Let X_t^I be the random variable indicating whether the incumbent had successfully innovated in the t^{th} year. Let X_t^E be the random variable indicating whether the entrant had successfully innovated in the t^{th} year. The random variables take the value of 1 with success and 0 with failure.

$$X_t^I = \begin{cases} 1 & \text{if R\&D succeeds in } t^{th} \text{ for } \mathbf{I} \\ 0 & \text{if R\&D fails in } t^{th} \text{ year for } \mathbf{I} \end{cases} \qquad X_t^E = \begin{cases} 1 & \text{if R\&D succeeds in } t^{th} \text{ year for } \mathbf{E} \\ 0 & \text{if R\&D fails in } t^{th} \text{ year for } \mathbf{E} \end{cases}$$

For simplicity, assume that previous R&D failures have **no impact** on current R&D. The firm neither learns from nor gets discouraged by earlier research failures. So X_t^I is statistically independent of X_s^I for all $s < t$. Similarly for X_t^E.

Consider the following conditional probability for the incumbent

$$P(X_t^I = x|\text{ no earlier R\&D success}) = \begin{cases} \theta_I & \text{if } x = 1 \\ 1 - \theta_I & \text{if } x = 0 \\ 0 & \text{otherwise} \end{cases}$$

and similarly for the entrant

$$P(X_t^E = x|\text{ no earlier R\&D success}) = \begin{cases} \theta_E & \text{if } x = 1 \\ 1 - \theta_E & \text{if } x = 0 \\ 0 & \text{otherwise} \end{cases}$$

$\{\theta_I, \theta_E\}$ are known as "transition probability" (for the incumbent and entrant, respectively). They denote the probability that the firm will "transition" from one state (here, R&D failure) to another (here, R&D success).

(a) Briefly explain why the following probability mass function statements are correct:

$$\begin{aligned} P(\{X_1^E = 0, X_2^E = 1\}) &= P(X_1^E = 0) \cdot P(X_2^E = 1 \mid X_1^E = 0) \\ P(\{X_1^E = 0, X_2^E = 0, X_3^E = 1\}) &= P(X_1^E = 0) \cdot P(X_2^E = 0 \mid X_1^E = 0) \cdot P(X_3^E = 1 \mid \{X_1^E = 0, X_2^E = 0\}) \end{aligned}$$

ANS: The first statement stems from the direct application of the definition of conditional probability: $P(A|B) = P(A, B)/P(B)$, therefore $P(A, B) = P(A|B)P(B)$.

The second statement stems from the iterated application of the definition of conditional probability. Note that $P(A|B, C) = P(A, B|C)/P(B|C)$ and $P(A, B|C) = P(A, B, C)/P(C)$; therefore $P(A, B, C) = P(A|B, C)\ P(B|C)\ P(C)$

(b) Let Y^E denote the random variable that indicates the number of years of research until the entrant achieves R&D success. So the event $Y^E = 5$ corresponds to $\{X_1^E = 0, X_2^E = 0, \ldots X_5^E = 1\}$ Explain why the probability mass function for Y^E is as follows:

$$P(Y^E = y) = \begin{cases} (1 - \theta_E)^{y-1}\,\theta_E & \text{if } y \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

ANS: The answer stems from the results shown in part **(b)**. Note that $\{Y^E = 1\}$ corresponds to $\{X_1^E = 1\}$, $\{Y^E = 2\}$ to $\{X_1^E = 0, X_E^2 = 1\}$, $\{Y^E = 3\}$ to $\{X_1^E = 0, X_2^E = 0, X_3^E = 1\}$, and so forth. So the $P(Y^E = 1) = P(\{X_1^E = 1\})$, $P(Y^E = 2) = P(\{X_1^E = 0, X_2^E = 1\})$, $P(Y^E = 3) = P(\{X_1^E = 0, X_2^E = 0, X_3^E = 1\})$, and so forth.

(c) Write down the cumulative distribution function for Y^E

ANS:

$$
\begin{aligned}
F_{Y^E}(y) &= P(Y^E \leq y) \\
&= \begin{cases} \sum_{s=1}^{y} P(Y^E = s) & \text{if } y \geq 1 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \sum_{s=1}^{y} (1 - \theta_E)^{s-1} \theta_E & \text{if } y \geq 1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

The above assumes that only integer values of y are allowed. If we allow non-integer reals, we would have to tweak the above to note that probability mass is only gained at each positive integer value.

You can further simplify the above power summation, but that is not required for this particular question.