Probability and Statistics

Exercise sheet 10

Exercise 10.1 Let X and Y be independent random variables such that $X \sim \text{Poi}(\lambda)$ and $Y \sim \text{Poi}(\mu)$. Show that given $X+Y=s$, the conditional distribution of X is $\text{Bin}\left(s, \frac{\lambda}{\lambda+\mu}\right)$.

Solution 10.1 We have $X \sim \text{Poi}(\lambda)$ and $Y \sim \text{Poi}(\mu)$ with $\lambda, \mu \in (0, +\infty)$, and X and Y are independent. For $s \in \mathbb{N}$,

$$P(X=k \mid X+Y=s) = \frac{P(X=k, X+Y=s)}{P(X+Y=s)} = \begin{cases} \frac{P(X=k)P(Y=s-k)}{P(X+Y=s)} & \text{if } 0 \leq k \leq s, \\ 0 & \text{otherwise.} \end{cases}$$

Since $X+Y \sim \text{Poi}(\lambda+\mu)$, it follows that

$$\begin{aligned} \frac{P(X=k)P(Y=s-k)}{P(X+Y=s)} &= \frac{e^{-\lambda}\lambda^k}{k!} \frac{e^{-\mu}\mu^{s-k}}{(s-k)!} \frac{s!}{e^{-(\lambda+\mu)}(\lambda+\mu)^s} \\ &= \frac{\lambda^k \mu^{s-k}}{(\lambda+\mu)^s} \binom{s}{k} \\ &= \left(\frac{\lambda}{\lambda+\mu}\right)^k \left(\frac{\mu}{\lambda+\mu}\right)^{s-k} \binom{s}{k}. \end{aligned}$$

Letting $p = \frac{\lambda}{\lambda+\mu}$,

$$P(X=k \mid X+Y=s) = \begin{cases} \binom{s}{k} p^k (1-p)^{s-k}, & k \in \{0, ..., s\} \\ 0, & \text{otherwise.} \end{cases}$$

In other words, $X \mid X+Y=s \sim \text{Bin}(s,p)$.

Exercise 10.2 Let X, Y and Z be $\overset{\text{iid}}{\sim} \text{Exp}(1)$.

(a) Find $E[X\sqrt{X+Y}]$.

(b) Find $P(X < 2Y < 3Z)$.

Solution 10.2

(a) Note that by the symmetry of the distribution (exchanging the roles of X and Y),

$$E(X\sqrt{X+Y}) = E(Y\sqrt{Y+X}).$$

Hence,

$$E(X\sqrt{X+Y}) = \frac{1}{2}E((X+Y)\sqrt{X+Y}) = \frac{1}{2}E((X+Y)^{\frac{3}{2}}).$$

Recall that if $X_1, ..., X_n$ are independent such that $X_i \sim G(\alpha_i, \beta)$, $\alpha_1, ..., \alpha_n$, $\beta > 0$, then $X_1 + ... + X_n \sim G(\alpha_1 + ... + \alpha_n, \beta)$.

Therefore, $X+Y \sim G(2,1)$ (since $\mathrm{Exp}(1) \stackrel{\mathrm{d}}{=} G(1,1)$).
It follows that

$$\begin{aligned}
E((X+Y)^{\frac{3}{2}}) &= E(S^{\frac{3}{2}}) \quad \text{(where } S \sim G(2,1)) \\
&= \int_0^\infty \frac{1^2}{\Gamma(2)} s e^{-s} s^{\frac{3}{2}} ds \\
&= \int_0^\infty s^{\frac{5}{2}} e^{-s} ds \quad \text{(since } \Gamma(2) = 1! = 1) \\
&= \Gamma\left(\frac{7}{2}\right) \int_0^\infty \frac{1^{\frac{7}{2}}}{\Gamma\left(\frac{7}{2}\right)} s^{\frac{7}{2}} e^{-s} ds \\
&= \Gamma\left(\frac{7}{2}\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
E(X\sqrt{X+Y}) &= \frac{1}{2}\Gamma\left(\frac{7}{2}\right) \\
&= \frac{1}{2}\Gamma\left(\frac{5}{2}+1\right) \\
&= \frac{1}{2}\frac{5}{2}\Gamma\left(\frac{5}{2}\right) \\
&= \ldots \\
&= \frac{1}{2}\frac{5}{2}\frac{3}{2}\frac{1}{2}\Gamma\left(\frac{1}{2}\right) \\
&= \frac{15}{16}\sqrt{\pi}.
\end{aligned}$$

(b)

$$P(X < 2Y < 3Z) = \iiint_D e^{-x} e^{-y} e^{-z} \, dx \, dy \, dz$$

with

$$D = \{(x,y,z) \in \mathbb{R}^3 : 0 \le x < 2y < 3z\} = \{(x,y,z) \in \mathbb{R}^3 : x \ge 0, y > \frac{x}{2}, z > \frac{2}{3}y\}.$$

By Fubini's Theorem, we have that

$$
\begin{aligned}
P(X < 2Y < 3Z) &= \int_0^\infty \int_{\frac{x}{2}}^\infty \int_{\frac{2}{3}y}^\infty e^{-z} e^{-y} e^{-x} dx dy dz \\
&= \int_0^\infty \int_{\frac{x}{2}}^\infty \left(\int_{\frac{2}{3}y}^\infty e^{-z} dz \right) e^{-y} e^{-x} dy dx \\
&= \int_0^\infty \int_{\frac{x}{2}}^\infty e^{-\frac{2}{3}y - y} e^{-x} dy dx \\
&= \int_0^\infty \left(\int_{\frac{x}{2}}^\infty e^{-\frac{5}{3}y} dy \right) e^{-x} dx \\
&= \int_0^\infty \frac{3}{5} e^{-\frac{5}{6}x - x} dx \\
&= \frac{3}{5} \int_0^\infty e^{-\frac{11}{6}x} dx \\
&= \frac{3}{5} \frac{6}{11} \\
&= \frac{18}{55}.
\end{aligned}
$$

Exercise 10.3 Suppose $X \sim \mathcal{N}(0,1)$ and $Y \mid X = x \sim \mathcal{N}(x+1,1)$.

(a) What is the marginal distribution of Y?

(b) Find $\text{cov}(X,Y)$ and the correlation of X and Y.

(c) Find the conditional distribution of X given $Y = y$.

Solution 10.3

(a) The marginal density of Y is given by

$$
f_Y(y) = \int_{\mathbb{R}} f(x,y) dx
$$

where f is the joint density of (X,Y). Now,

$$
\begin{aligned}
f(x,y) &= f(y \mid x) f_X(x) \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x-1)^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \\
&= \frac{1}{2\pi} e^{-\frac{1}{2}((y-1)^2 - 2(y-1)x + x^2 + x^2)} \\
&= \frac{1}{2\pi} e^{-\frac{1}{2}(2x^2 - 2(y-1)x + (y-1)^2)} \\
&= \frac{1}{2\pi} e^{-(x^2 - (y-1)x + \frac{(y-1)^2}{2})} \\
&= \frac{1}{2\pi} e^{-(x^2 - 2\frac{(y-1)}{2}x + \frac{(y-1)^2}{4} + \frac{(y-1)^2}{4})} \\
&= \frac{1}{2\pi} e^{(-\frac{(y-1)^2}{4})} e^{-(x - \frac{y-1}{2})^2}.
\end{aligned}
$$

$$
\begin{aligned}
f_Y(y) &= \frac{1}{2\pi} e^{-\frac{(y-1)^2}{4}} \int_{\mathbb{R}} e^{-(x-\frac{y-1}{2})^2} dx \\
&= \frac{1}{\sqrt{2\pi}} \frac{e^{-\frac{(y-1)^2}{4}}}{\sqrt{2}} \left(\frac{1}{\sqrt{2\pi}} \left(\frac{1}{\sqrt{2}}\right)^{-1} \int_{\mathbb{R}} e^{-\frac{\left(x-\frac{y-1}{2}\right)^2}{2\frac{1}{2}}} dx \right) \\
&= \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-\frac{(y-1)^2}{2\times 2}}
\end{aligned}
$$

(since we take the integral of the density of $\mathcal{N}\left(\frac{y-1}{2}, \frac{1}{2}\right)$) and so $Y \sim \mathcal{N}(1, 2)$.

(b)

$$
\begin{aligned}
\operatorname{cov}(X,Y) &= E(XY) - E(X)E(Y) \\
&= E(XY) && \text{(as } E(X) = 0) \\
&= E(E(XY \mid X)) && \text{(iterated expectation)} \\
&= E(XE(Y \mid X)) \\
&= E(X(X+1)) \\
&= E(X^2) + E(X) \\
&= E(X^2) && \text{(as } E(X) = 0) \\
&= \operatorname{var}(X) && \text{(as } E(X) = 0) \\
&= 1
\end{aligned}
$$

and

$$
\begin{aligned}
\operatorname{corr}(X,Y) &= \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var}(X)}\sqrt{\operatorname{var}(Y)}} \\
&= \frac{1}{\sqrt{1}\sqrt{2}} \\
&= \frac{1}{\sqrt{2}}.
\end{aligned}
$$

(c) Note that the previous calculations give already that

$$
\begin{aligned}
f(x,y) &= \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-\frac{(y-1)^2}{4}} \frac{1}{\sqrt{2\pi}} \sqrt{2} e^{-\frac{(x-\frac{y-1}{2})^2}{2\frac{1}{2}}} \\
&= f_Y(y) f(x \mid y)
\end{aligned}
$$

implying that $X \mid Y = y \sim \mathcal{N}\left(\frac{y-1}{2}, \frac{1}{2}\right)$.

Exercise 10.4 Gaussian (normal) vectors.

A vector $X = (X_1, ..., X_n)^T$ is said to be a Gaussian vector if there is a matrix $A \in \mathbb{R}^{n\times n}$, a vector $Z = (Z_1, ..., Z_n)^T$ with $Z_1, ..., Z_n \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$ and $\mu \in \mathbb{R}^n$ such that

$$
X \overset{\text{d}}{=} \mu + AZ. \tag{1}
$$

In this case, note that $E(X) = \mu$ and $\operatorname{var}(X) = AA^T =: \Sigma$. Here, the covariance matrix Σ is not necessarily invertible. If it is, then X admits a density with respect to Lebesgue measure on $\mathbb{R}^n$.

(a) Show that if X is a Gaussian vector, then any linear combination of $X_1, ..., X_n$ is a normal random variable.

(b) We want to show that the condition in (a) is also sufficient: i.e., that if any linear combination of $X_1, ..., X_n$ is a normal random variable, then X is a Gaussian vector.

1. Explain why you can write $\Sigma = P^T D P$ with P orthogonal and D a diagonal matrix with entries $\lambda_1, ..., \lambda_n \geq 0$. Let $A := P^T D^{\frac{1}{2}}$, where $D^{\frac{1}{2}}$ is defined as the diagonal matrix with entries $\lambda_1^{\frac{1}{2}}, ..., \lambda_n^{\frac{1}{2}}$.
2. For a fixed $v \in \mathbb{R}^n$, find the distribution of $v^T(\mu + AZ)$ with $Z = (Z_1, ..., Z_n)^T$ with $Z_1, ..., Z_n \overset{\text{iid}}{\sim} \mathcal{N}(0, 1)$.
   Hint: From a previous exercise sheet, you know that a linear combination of i.i.d standard Gaussians has a Gaussian distribution $\mathcal{N}(a, b^2)$. What are the parameters a, b^2 here?
3. What is the distribution of $v^T X$?
4. To conclude, use that fact that two random vectors W_1, W_2 in $\mathbb{R}^n$ have the same distribution if and only if

$$v^T W_1 \overset{\text{d}}{=} v^T W_2 \quad \forall v \in \mathbb{R}^n.$$

(c) Let $X \sim \mathcal{N}(0, 1)$ and Z be a discrete random variable such that $X \perp\!\!\!\perp Z$ and $P(Z = -1) = P(Z = 1) = \frac{1}{2}$.

Consider the random variable Y, defined by

$$Y = \begin{cases} X, & \text{if } Z = 1 \\ -X, & \text{if } Z = -1. \end{cases}$$

- Show that marginally, $Y \sim \mathcal{N}(0, 1)$.
- Find $P(X + Y = 0)$.
- Is (X, Y) a Gaussian vector? What do you conclude from this exercise?

Solution 10.4

(a) For this question, we recall the following facts:

- If $X_1, ..., X_n$ are independent with $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, then

$$\sum_{i=1}^n a_i X_i \sim \mathcal{N}\left(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2\right).$$

- If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $X + c \sim \mathcal{N}(\mu + c, \sigma^2)$ (the latter can be shown easily and is left as a small exercise).

Let

$$a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Then

$$\begin{aligned}
\sum_{i=1}^{n} a_i X_i &= a^T X \\
&\overset{\mathrm{d}}{=} a^T \mu + a^T A Z \\
&= c + b^T Z \qquad \text{(where } c = a^T \mu \text{ and } b = A^T a) \\
&= c + \sum_{i=1}^{n} b_i Z_i \\
&= \mathcal{N}\left(c, \sum_{i=1}^{n} b_i^2\right)
\end{aligned}$$

using the two known facts recalled above.

(b) 1. By definition of Σ, we know that

$$\Sigma_{ij} = \mathrm{cov}(X_i, X_j), \quad 1 \le i, j \le n.$$

For any vector $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$,

$$\begin{aligned}
v^T \Sigma v &= \sum_{1 \le i,j \le n} v_i v_j \Sigma_{ij} \\
&= \sum_{1 \le i,j \le n} v_i v_j \mathrm{cov}(X_i, X_j) \\
&= \mathrm{var}\left(\sum_{i=1}^{n} \sigma_i X_i\right) \\
&\ge 0.
\end{aligned}$$

Thus, Σ is symmetric and also positive semi-definite. Now, symmetry implies that Σ is diagonalisable and we can find an orthogonal matrix P in $\mathbb{R}^{n \times n}$, and D a diagonalisable matrix with diagonal entries $\lambda_1, ..., \lambda_n$ such that

$$\Sigma = P^T D P.$$

Furthermore, $\lambda_i \ge 0\ \forall i \in \{1, ..., n\}$, since if we take $v = P^{-1} e_i$ (e_i is the i^{th} standard basis vector in $\mathbb{R}^n$), then

$$0 \le v^T \Sigma v = e_i^T D e_i = \lambda_i.$$

Then, define $D^{\frac{1}{2}}$ as the diagonal matrix with entries $\sqrt{\lambda_1}, ..., \sqrt{\lambda_n}$ (now we know we can define it).

2. Using the same arguments as for (a), for $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$, we have that

$$v^T(\mu + AZ) \sim \mathcal{N}\left(v^T \mu, \sum_{i=1}^{n} b_i^2\right)$$

where $b = A^T v$.

Now, note that

$$\begin{aligned}\sum_{i=1}^n b_i^2 &= b^T b \\ &= v^T A A^T b \\ &= v^T P^T D^{\frac{1}{2}} D^{\frac{1}{2}} P v \\ &= v^T P^T D P v \\ &= v^T \Sigma v\end{aligned}$$

and therefore, $v^T(\mu + AZ) \sim \mathcal{N}(v^T\mu, v^T\Sigma v)$.

3. Since we assume now that any linear combination of $X_1, ..., X_n$ is a normal random variable, it follows that $v^T X = v_1 X_1 + ... + v_n X_n$ is normal. To characterise its distribution, it is enough to compute its mean and variance.

$$E(v^T X) = v^T E(X) = v^T \mu,$$

$$\begin{aligned}\mathrm{var}(v^T X) &= \mathrm{var}\left(\sum_{i=1}^n v_i X_i\right) \\ &= \sum_{1 \le i,j \le n} v_i v_j \mathrm{cov}(X_i, X_j) \\ &= \sum_{1 \le i,j \le n} v_i v_j \Sigma_{i,j} \\ &= v^T \Sigma v.\end{aligned}$$

so that $v^T X \sim \mathcal{N}(v^T\mu, v^T\Sigma v)$.

4. We have $\forall v \in \mathbb{R}^n$, $v^T X \overset{\mathrm{d}}{=} v^T(\mu + AZ)$ with $A = P^T D^{\frac{1}{2}}$. We conclude from the hint that

$$X \overset{\mathrm{d}}{=} \mu + AZ.$$

(c) • To find the distribution of Y, it is enough to compute its cdf. For $y \in \mathbb{R}$:

$$\begin{aligned}F_Y(y) &= P(Y \le y) \\ &= P(\{X \le y, Z = 1\} \cup \{-X \le y, Z = -1\}) \\ &= P(X \le y, Z = 1) + P(-X \le y, Z = -1) \\ &= P(X \le y)P(Z = 1) + P(-X \le y)P(Z = -1) \\ &= \frac{1}{2}\Phi(y) + \frac{1}{2}(1 - \Phi(-y)) \qquad \text{(since } \mathcal{N}(0,1) \text{ is continuous)} \\ &= \frac{1}{2}\Phi(y) + \frac{1}{2}\Phi(y) \qquad \text{(by symmetry of the standard normal)} \\ &= \Phi(y).\end{aligned}$$

Thus, $Y \sim \mathcal{N}(0, 1)$.

- $$\begin{aligned} P(X+Y=0) &= P(2X=0, Z=1) + P(X-X=0, Z=-1) \\ &= P(X=0)P(Z=1) + P(Z=-1) \\ &= \frac{1}{2}. \end{aligned}$$
- If (X,Y) were a Gaussian vector, then any linear combination of X and Y is a normal random variable, and hence should either be absolutely continuous or a constant, neither of which is the case, since we would need either $P(X+Y=0)=0$ or $P(X+Y=0)=1$. Thus (X,Y) is not a Gaussian vector.
- From this exercise we can conclude that putting together two random variables (on the same probability space), which have normal marginal distributions, does not guarantee that they form a Gaussian vector.

Exercise 10.5 (optional).

The goal of this exercise is to manipulate the Jacobian formula to obtain the density of a convolution.

Let X and Y be two independent random variables with density f_X and f_Y respectively. We are interested in deriving the density of the random variable $X+Y$.

(a) Let $S := X+Y$, $T := Y$ and consider the map

$$\begin{aligned} g : \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x,y) &\mapsto (s,t) = (x+y, y). \end{aligned}$$

Show that g is bijective with Jacobian $\neq 0$ at any $(x,y) \in \mathbb{R}^2$. Conclude that the random pair (S,T) has point density $f_{(S,T)}$ given by

$$f_{(S,T)}(s,t) = f_X(s-t)f_Y(t).$$

(b) Conclude that

$$f_S(s) = \int_{\mathbb{R}} f_X(s-y)f_Y(y)dy.$$

(c) Give the density of the convolution $X+Y$ in the following cases:

- $X \sim \text{Exp}(\lambda)$ and $Y \sim \text{Exp}(\mu)$. Do you recognise the distribution when $\mu = \lambda$?
- $X \sim G(\alpha_1, \beta)$ and $G(\alpha_2, \beta)$.
- $X \sim \mathcal{N}(0, \sigma_1^2)$ and $Y \sim \mathcal{N}(0, \sigma_2^2)$.

Solution 10.5

(a)
$$g(x,y) = (x+y,y) = (s,t) \Leftrightarrow x = s-t, y = t.$$

Thus, g is bijective and $g^{-1}(s,t) = (s-t,t)$. Also, g is differentiable and

$$\nabla g(x,y) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Thus the Jacobian is given by

$$J_g(x,y) = \det(\nabla g) = 1.$$

(b) Using the Jacobian formula, it follows that

$$f_{(S,T)}(s,t) = f_{(X,Y)}(s-t,t) \times \frac{1}{1} = f_X(s-t)f_Y(t)$$

by independence of X and Y.

Hence

$$f_S(s) = \int_{\mathbb{R}} f_X(s-t)f_Y(t)dt = \int_{\mathbb{R}} f_X(s-y)f_Y(y)dy.$$

(c)
- We compute:

$$\begin{aligned} f_S(s) &= \int_{\mathbb{R}} \lambda e^{-\lambda(s-y)} \mathbb{1}_{0\le s-y} \mu e^{-\mu y} \mathbb{1}_{y\ge 0}\, dy \\ &= \lambda\mu e^{-\lambda s}\mathbb{1}_{s\ge 0} \int_0^s e^{(\lambda-\mu)y}\, dy \end{aligned}$$

  (rewriting the conditions as $\mathbb{1}_{0\le s-y}\mathbb{1}_{y\ge 0} = \mathbb{1}_{s\ge 0}\mathbb{1}_{0\le y\le s}$).
  Hence,

$$\begin{aligned} f_S(s) &= \begin{cases} \lambda\mu e^{-\lambda s}\mathbb{1}_{s\ge 0}\frac{e^{(\lambda-\mu)s}-1}{\lambda-\mu} & \text{if } \lambda \neq \mu \\ \lambda^2 e^{-\lambda s} s \mathbb{1}_{s\ge 0} & \text{if } \lambda = \mu \end{cases} \\ &= \begin{cases} \frac{\lambda\mu}{\lambda-\mu}(e^{-\mu s} - e^{-\lambda s})\mathbb{1}_{s\ge 0} & \text{if } \lambda \neq \mu \\ \lambda^2 s e^{-\lambda s}\mathbb{1}_{s\ge 0} & \text{if } \lambda = \mu. \end{cases} \end{aligned}$$

  Once again, we find that $X+Y \sim G(1,2)$ if X, Y are i.i.d $\text{Exp}(\lambda)$.
- For $X \sim G(\alpha_1,\beta)$ and $Y \sim G(\alpha_2,\beta)$,

$$\begin{aligned} f_S(s) &= \mathbb{1}_{s\ge 0}\int_0^s \frac{\beta^{\alpha_1}}{\Gamma(\alpha_1)}(s-y)^{\alpha_1-1}e^{-\beta(s-y)}\frac{\beta^{\alpha_2}}{\Gamma(\alpha_2)}y^{\alpha_2-1}e^{-\beta y}\, dy \\ &= \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}e^{-\beta s}\mathbb{1}_{s\ge 0}\int_0^s (s-y)^{\alpha_1-1}y^{\alpha_2-1}\, dy. \end{aligned}$$

  Putting $t = \frac{y}{s}$ so that $dt = \frac{dy}{s}$,

$$\begin{aligned} \int_0^s (s-y)^{\alpha_1-1}y^{\alpha_2-1}dy &= \int_0^s s(1-t)^{\alpha_1-1}t^{\alpha_2-1}s^{\alpha_1+\alpha_2-2}dt \\ &= s^{\alpha_1+\alpha_2-1}\int_0^1 (1-t)^{\alpha_1}t^{\alpha_2-1}dt \\ &= s^{\alpha_1+\alpha_2-1}\frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)} \end{aligned}$$

  (since $\frac{\Gamma(\alpha_1)+\Gamma(\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}t^{\alpha_1-1}(1-t)^{\alpha_2-2}\mathbb{1}_{t\in(0,1)}$ is the density of $\text{Beta}(\alpha_1,\alpha_2)$).
  Hence,

$$f_S(s) = \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)}s^{\alpha_1+\alpha_2-1}e^{-\beta s}\mathbb{1}_{s\ge 0}.$$

  We find again that $X+Y \sim G(\alpha_1+\alpha_2,\beta)$.

- For $X \sim \mathcal{N}(0,\sigma_1^2)$ and $Y \sim \mathcal{N}(0,\sigma_2^2)$,

$$f_S(s) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(s-y)^2}{2\sigma_2^2}} \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{y^2}{2\sigma_2^2}}\, dy$$

We manipulate the exponent to make this easier to integrate:

$$\begin{aligned}
\frac{(s-y)^2}{2\sigma_1^2} + \frac{y^2}{2\sigma_2^2} &= \frac{1}{2}\frac{\sigma_2^2(s^2-2sy+y^2)+\sigma_1^2 y^2}{\sigma_1^2\sigma_2^2} \\
&= \frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}\left(y^2 - \frac{2\sigma_2^2}{\sigma_1^2+\sigma_2^2} sy + \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2} s^2\right) \\
&= \frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}\left(\left(y - \frac{\sigma_2^2 s}{\sigma_1^2+\sigma_2^2}\right)^2 - \frac{\sigma_2^4}{(\sigma_1^2+\sigma_2^2)^2} s^2 + \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2} s^2\right) \\
&= \frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}\left(y - \frac{\sigma_2^2 s}{\sigma_1^2+\sigma_2^2}\right)^2 + \frac{s^2}{2(\sigma_1^2+\sigma_2^2)}.
\end{aligned}$$

Thus, letting $\sigma^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}$, we get that

$$\begin{aligned}
f_S(s) &= \frac{1}{\sqrt{2\pi}\sigma_1}\frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{s^2}{2(\sigma_1^2+\sigma_2^2)}} \int_{\mathbb{R}} e^{-\frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}\left(y-\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)^2} dy \\
&= \frac{1}{\sqrt{2\pi}\sigma_1\sigma_2}\sigma e^{-\frac{s^2}{2(\sigma_1^2+\sigma_2^2)}} \frac{1}{\sqrt{2\pi}\sigma}\int_{\mathbb{R}} e^{-\frac{\left(y-\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}s\right)^2}{2\sigma^2}} dy \\
&= \frac{1}{\sqrt{2\pi}}\frac{\sigma_1\sigma_2}{\sigma_1\sigma_2\sqrt{\sigma_1^2+\sigma_2^2}} e^{-\frac{s^2}{2(\sigma_1^2+\sigma_2^2)}} \\
&= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\sigma_1^2+\sigma_2^2}} e^{-\frac{s^2}{2(\sigma_1^2+\sigma_2^2)}}.
\end{aligned}$$

Hence we find the expected result that

$$X + Y \sim \mathcal{N}(0, \sigma_1^2+\sigma_2^2).$$