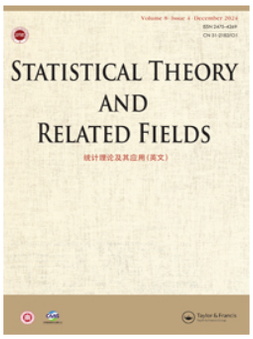

Statistical Theory and Related Fields

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tstf20

A note on the probability structure of a doubly randomized delayed start design

Zhao Yang

To cite this article: Zhao Yang (31 Jan 2025): A note on the probability structure of a doubly randomized delayed start design, Statistical Theory and Related Fields, DOI: 10.1080/24754269.2025.2457279

To link to this article: https://doi.org/10.1080/24754269.2025.2457279

STATISTICAL THEORY AND RELATED FIELDS
https://doi.org/10.1080/24754269.2025.2457279

Taylor & Francis
Taylor & Francis Group

A note on the probability structure of a doubly randomized delayed start design

Zhao Yang

Keros Therapeutics Inc. , Lexington, MA, USA

ABSTRACT

The innovative doubly randomized delayed start (DRDS) design has been implemented to tackle the well-known challenge of a high placebo response rate in clinical trials. This design begins with a conventional parallel design phase (period 1), followed by a subsequent phase (period 2) where subjects initially assigned to placebo and who did not respond are re-randomized to either the test drug or placebo. Chi, G. Y., Li, Y., Liu, Y., Lewin, D., & Lim, P. (2016 On clinical trials with a high placebo response rate. *Contemporary Clinical Trials Communications*, *2*, 34–53. https://doi.org/10.1016/j.conctc.2015.10.002) introduced a new statistical methodology with a conditional probability structure to account for the specific characteristics of the DRDS design. However, some critical formulas in Chi et al. (2016, p. 38) for this probability structure are incorrect. Here, we correct these formulas and provide a comprehensive technical background on deriving the probability structure for a DRDS design to support these corrections.

ARTICLE HISTORY
Received 21 September 2024
Accepted 18 January 2025

KEYWORDS
Doubly randomized delayed start; sequential parallel comparison design; placebo effect; truncation

1. Introduction

A high placebo response rate, which contributes to the high failure rate of trials, is a significant and well-documented challenge frequently encountered in clinical trials, especially in fields such as neurology, psychiatry, and pain management. Hegerl and Mergl (2010) provided an intuitive visual representation to elucidate the mechanism of high placebo response rates. A novel sequential parallel comparison design (SPCD) framework, which was aimed at increasing the efficiency of placebo-controlled psychiatric clinical trials, was proposed by Fava et al. (2003) to address this issue of high placebo response rate. The approach involves an initial standard parallel design period (i.e., period 1), followed by a second period (i.e., period 2) where subjects who were originally randomized to placebo and did not respond are re-randomized to either the test drug or placebo. SPCD is also sometimes referred to as the doubly randomized delayed start (DRDS) design (Liu et al., 2012). In this short communication, we use the term DRDS to align with the practice from Chi et al. (2016).

The DRDS design has been recognized by regulatory agencies as an innovative approach. However, these agencies have raised concerns regarding the various proposed statistical analysis methods and the clinical interpretability of the results. Chi et al. (2016) provided a comprehensive summary of these key issues and introduced a new statistical methodology that differs from the existing methods. Specifically, the currently available methods treat the second period of a DRDS design as an independent trial from period 1. In contrast, the proposal by Chi et al. (2016) incorporates a conditional probability structure for period 2, reflecting that subjects in period 2 are placebo non-responders from period 1 who were re-randomized to either the test drug or placebo. It is noted that some of the critical formulas for the probability structure in Chi et al. (2016, p. 38) are inaccurate. In this short communication, we provide corrections to these formulas and offer a general technical background for deriving the probability structure of a DRDS design to support these corrections.

The structure of this short communication is as follows. Section 2 introduces the technical notations used to present the probability structure of a DRDS design. Section 3 presents the results for the probability structure. The Appendix provides the general technical background for deriving the probability structure of a DRDS design.

2. General notations

Consider a trial employing a DRDS design, as illustrated in Figure 1. At the beginning of period 1, N subjects are randomly allocated to either the test (t) group or the placebo (p) group, with $n_{1,t}$ subjects assigned to the test group and $n_{1,p}$ subjects assigned to the placebo group. Let $Y_{1,t}$ and $Y_{1,p}$ represent the continuous clinical response variables of interest under the test group and the placebo group, respectively. Both are normally distributed, with

CONTACT Zhao Yang tonyyangsxz@gmail.com Keros Therapeutics Inc., 1050 Waltham St, Suite 302, Lexington, MA 02421, USA

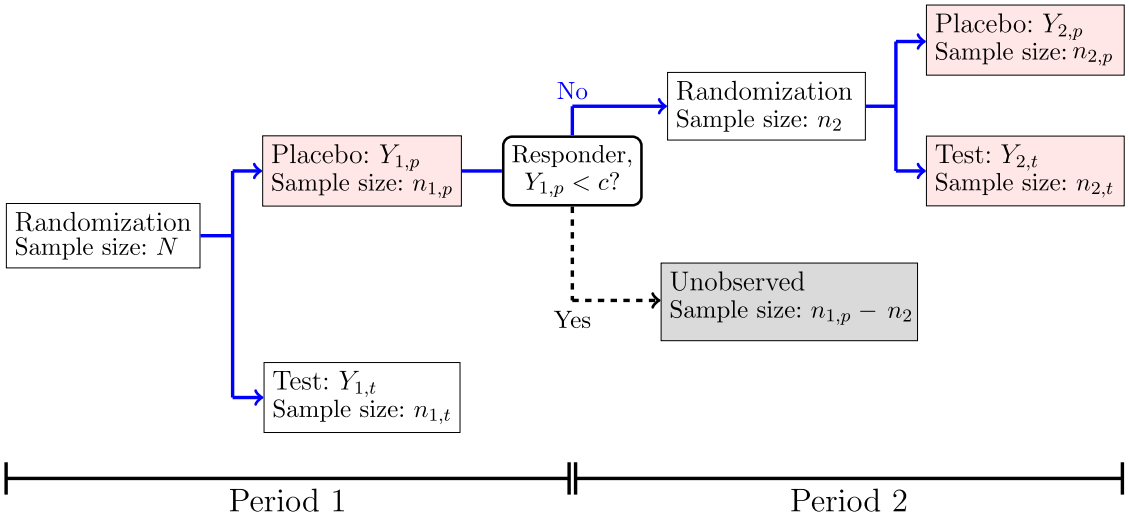

Figure 1. Basic study flow based on a DRDS design and its observed variables for each treatment group in each period.

$Y_{1,t} \sim N(\mu_{1,t}, \sigma^2_{1,t})$ and $Y_{1,p} \sim N(\mu_{1,p}, \sigma^2_{1,p})$. At the end of period 1, a pre-specified criterion will be applied to determine the response status of subjects in the placebo group who completed the period. Specifically, those in the placebo group identified as responders – i.e., if $Y_{1,p} \geq c$ – along with those who discontinued period 1 early, will be excluded from the second period of the study. In contrast, n_2 subjects in the placebo group who are classified as non-responders–i.e., $Y_{1,p} < c$–will be re-randomized to either the test or placebo group at the beginning of period 2, with $n_{2,t}$ subjects in the test group and $n_{2,p}$ subjects in the placebo group.

Suppose that all subjects in the placebo group at the end of period 1 were re-randomized in period 2 to the test group in period 2. Let $Y_{2,t}$ represent the clinical response variables of interest which follows a normal distribution, $Y_{2,t} \sim N(\mu_{2,t}, \sigma^2_{2,t})$. In this case, the pair of variables $(Y_{1,p}, Y_{2,t})$ follows a bivariate normal distribution with a correlation of $\rho_t = \text{Corr}(Y_{1,p}, Y_{2,t})$. Similarly, if all subjects in the placebo group at the end of period 1 were re-randomized to the placebo group in period 2, let $Y_{2,p}$ denote the clinical response variables of interest, which follows a normal distribution, $Y_{2,p} \sim N(\mu_{2,p}, \sigma^2_{2,p})$. In this scenario, the pair of variables $(Y_{1,p}, Y_{2,p})$ follows a bivariate normal distribution with a correlation of $\rho_p = \text{Corr}(Y_{1,p}, Y_{2,p})$. The bivariate normal distributions for the pairs of $(Y_{1,p}, Y_{2,t})$ and $(Y_{1,p}, Y_{2,p})$ are represented as follows:

$$(Y_{1,p}, Y_{2,t}) \sim N\left(\begin{pmatrix}\mu_{1,p}\\ \mu_{2,t}\end{pmatrix}, \begin{pmatrix}\sigma^2_{1,p} & \rho_t\sigma_{1,p}\sigma_{2,t}\\ \rho_t\sigma_{1,p}\sigma_{2,t} & \sigma^2_{2,t}\end{pmatrix}\right),$$
$$(Y_{1,p}, Y_{2,p}) \sim N\left(\begin{pmatrix}\mu_{1,p}\\ \mu_{2,p}\end{pmatrix}, \begin{pmatrix}\sigma^2_{1,p} & \rho_p\sigma_{1,p}\sigma_{2,p}\\ \rho_p\sigma_{1,p}\sigma_{2,p} & \sigma^2_{2,p}\end{pmatrix}\right).$$

However, according to the DRDS design, the subjects observed during period 2 are those who were assigned to the placebo group in period 1 and had an outcome value below a pre-defined threshold c, i.e., non-responder defined as $Y_{1,p} < c$. As a result, the variables observed in period 2 are not $Y_{2,t}$ or $Y_{2,p}$, but rather $Y_{2,t}|Y_{1,p} < c$ or $Y_{2,p}|Y_{1,p} < c$, which aligns with the framework of using singly truncated bivariate normal distributions. We will present its probability structure using this framework in the next section.

3. Probability structure of a DRDS design

For the pair with right truncation at $Y_{1,p} < c$, $(Y_{1,p}, Y_{2,t}|Y_{1,p} < c)$, its singly truncated bivariate normal distribution can be expressed as follows:

$$\begin{aligned}&(Y_{1,p}|Y_{1,p} < c, Y_{2,t}|Y_{1,p} < c)\\ &\quad\sim N\left(\begin{pmatrix}\mu_{1,p|Y_{1,p}<c}\\ \mu_{2,t|Y_{1,p}<c}\end{pmatrix}, \begin{pmatrix}\sigma^2_{1,p|Y_{1,p}<c} & \text{Cov}(Y_{1,p}, Y_{2,t}|Y_{1,p} < c)\\ \text{Cov}(Y_{1,p}, Y_{2,t}|Y_{1,p} < c) & \sigma^2_{2,t|Y_{1,p}<c}\end{pmatrix}\right).\end{aligned} \tag{1}$$

Similarly, for the pair with right truncation at $Y_{1,p} < c$, $(Y_{1,p}, Y_{2,p}|Y_{1,p} < c)$, its singly truncated bivariate normal distribution is

$$(Y_{1,p}|Y_{1,p} < c, Y_{2,p}|Y_{1,p} < c) \sim N\left(\begin{pmatrix}\mu_{1,p|Y_{1,p}<c}\\ \mu_{2,p|Y_{1,p}<c}\end{pmatrix}, \begin{pmatrix}\sigma^2_{1,p|Y_{1,p}<c} & \mathrm{Cov}(Y_{1,p}, Y_{2,p}|Y_{1,p} < c)\\ \mathrm{Cov}(Y_{1,p}, Y_{2,p}|Y_{1,p} < c) & \sigma^2_{2,p|Y_{1,p}<c}\end{pmatrix}\right), \tag{2}$$

where $\mathrm{Cov}(Y_{1,p}, Y_{2,t}|Y_{1,p} < c) = \rho_{t|Y_{1,p}<c}\sigma_{1,p|Y_{1,p}<c}\sigma_{2,t|Y_{1,p}<c}$, and $\mathrm{Cov}(Y_{1,p}, Y_{2,p}|Y_{1,p} < c) = \rho_{p|Y_{1,p}<c}\sigma_{1,p|Y_{1,p}<c}\sigma_{2,p|Y_{1,p}<c}$ in Equations (1) and (2), respectively. Defining $\tau = (c - \mu_{1,p})/\sigma_{1,p}$, the elements of interest in Equations (1) and (2) are given by

$$\begin{aligned}
\mu_{1,p|Y_{1,p}<c} &= \mu_{1,p} - \sigma_{1,p}\frac{\phi(\tau)}{\Phi(\tau)},\\
\mu_{2,t|Y_{1,p}<c} &= \mu_{2,t} - \rho_t\sigma_{2,t}\frac{\phi(\tau)}{\Phi(\tau)},\\
\mu_{2,p|Y_{1,p}<c} &= \mu_{2,p} - \rho_p\sigma_{2,p}\frac{\phi(\tau)}{\Phi(\tau)},\\
\sigma^2_{1,p|Y_{1,p}<c} &= \sigma^2_{1,p}\left[1 - \frac{\tau\phi(\tau)}{\Phi(\tau)} - \left(\frac{\phi(\tau)}{\Phi(\tau)}\right)^2\right],\\
\sigma^2_{2,t|Y_{1,p}<c} &= \sigma^2_{2,t}\left[1 - \rho_t^2\frac{\tau\phi(\tau)}{\Phi(\tau)} - \rho_t^2\left(\frac{\phi(\tau)}{\Phi(\tau)}\right)^2\right],\\
\sigma^2_{2,p|Y_{1,p}<c} &= \sigma^2_{2,p}\left[1 - \rho_p^2\frac{\tau\phi(\tau)}{\Phi(\tau)} - \rho_p^2\left(\frac{\phi(\tau)}{\Phi(\tau)}\right)^2\right],\\
\mathrm{Cov}(Y_{1,p}, Y_{2,t}|Y_{1,p} < c) &= \rho_t\sigma_{1,p}\sigma_{2,t}\left[1 - \frac{\tau\phi(\tau)}{\Phi(\tau)} - \left(\frac{\phi(\tau)}{\Phi(\tau)}\right)^2\right],\\
\mathrm{Cov}(Y_{1,p}, Y_{2,p}|Y_{1,p} < c) &= \rho_p\sigma_{1,p}\sigma_{2,p}\left[1 - \frac{\tau\phi(\tau)}{\Phi(\tau)} - \left(\frac{\phi(\tau)}{\Phi(\tau)}\right)^2\right],\\
\rho_{t|Y_{1,p}<c} &= \rho_t\sqrt{\left\{\rho_t^2 + (1-\rho_t^2)\left[1 - \frac{\tau\phi(\tau)}{\Phi(\tau)} - \left(\frac{\phi(\tau)}{\Phi(\tau)}\right)^2\right]^{-1}\right\}^{-1}},\\
\rho_{p|Y_{1,p}<c} &= \rho_p\sqrt{\left\{\rho_p^2 + (1-\rho_p^2)\left[1 - \frac{\tau\phi(\tau)}{\Phi(\tau)} - \left(\frac{\phi(\tau)}{\Phi(\tau)}\right)^2\right]^{-1}\right\}^{-1}}.
\end{aligned}$$

All of these identities can be readily derived from the general technical results presented in the Appendix. With the underlying conditional probability structure for a DRDS design as described above, the adjusted treatment effect estimation and its hypothesis testing can proceed based on the methods proposed by Chi et al. (2016).

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Zhao Yang http://orcid.org/0009-0008-6011-6802

References

Chi, G. Y., Li, Y., Liu, Y., Lewin, D., & Lim, P. (2016). On clinical trials with a high placebo response rate. *Contemporary Clinical Trials Communications*, *2*, 34–53. https://doi.org/10.1016/j.conctc.2015.10.002

Fava, M., Evins, A. E., Dorer, D. J., & Schoenfeld, D. A. (2003). The problem of the placebo response in clinical trials for psychiatric disorders: Culprits, possible remedies, and a novel study design approach. *Psychotherapy and Psychosomatics*, *72*(3), 115–127. https://doi.org/10.1159/000069738

Hegerl, U., & Mergl, R. (2010). The clinical significance of antidepressant treatment effects cannot be derived from placebo-verum response differences. *Journal of Psychopharmacology*, *24*(4), 445–448. https://doi.org/10.1177/0269881109106930
Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). *Continuous Univariate Distributions* (2nd ed., Vol. 1). John Wiley & Sons.
Kotz, S., Balakrishnan, N., & Johnson, N. L. (2000). *Continuous Multivariate Distributions, Volume 1: Models and Applications* (2nd ed.). John Wiley & Sons.
Liu, Q., Lim, P., Singh, J., Lewin, D., Schwab, B., & Kent, J. (2012). Doubly randomized delayed-start design for enrichment studies with responders or nonresponders. *Journal of Biopharmaceutical Statistics*, *22*(4), 737–757. https://doi.org/10.1080/10543406.2012.678234

Appendix

This appendix offers a general technical background for deriving the probability structure of a DRDS design, using notation that differs slightly from the main text. Additionally, since single truncation is a specific instance of double truncation, the technical background begins with the broader context of double truncation, with the single truncation case addressed appropriately at the end of this section.

Given a univariate random variable X that follows a standard normal distribution, the probability density function (pdf) and cumulative distribution function (cdf) are represented by $\phi(x)$ and $\Phi(x)$, respectively. It is known that $\phi'(x) = -x\phi(x)$ and $\Phi'(x) = \phi(x)$. If we define $Y_1 = \mu_1 + \sigma_1 X$, then Y_1 follows a normal distribution with mean μ_1 and variance σ_1^2, i.e., $Y_1 \sim N(\mu_1, \sigma_1^2)$. Now, consider the doubly truncated normal distribution where Y_1 is restricted to the interval $\mathcal{A} = [a, b]$ with $-\infty < a < b < \infty$. The probability of Y_1 lying within $\mathcal{A}$ is given by $\Phi((b-\mu_1)/\sigma_1) - \Phi((a-\mu_1)/\sigma_1)$. The pdf of the resulting truncated distribution is then expressed as $f(y_1; \mu_1, \sigma_1 | a \le y_1 \le b) = [\sigma_1^{-1}\phi((y_1-\mu_1)/\sigma_1)]/[\Phi((b-\mu_1)/\sigma_1) - \Phi((a-\mu_1)/\sigma_1)]$, the moment generating function (mgf) for this truncated distribution can then be derived as follows:

$$M_{Y_1}(t) = \mathrm{E}\left(e^{tY_1} | Y_1 \in \mathcal{A}\right) = \mathsf{R}\exp\left(\mu_1 t + \frac{t^2\sigma_1^2}{2}\right),$$

where

$$\mathsf{R} = \left[\Phi\left(\frac{b-\widetilde{\mu}}{\sigma_1}\right) - \Phi\left(\frac{a-\widetilde{\mu}}{\sigma_1}\right)\right]\left[\Phi(\tau_b) - \Phi(\tau_a)\right]^{-1},$$
$$\widetilde{\mu} = \mu_1 + t\sigma_1^2, \quad \tau_b = \frac{b-\mu_1}{\sigma_1}, \quad \tau_a = \frac{a-\mu_1}{\sigma_1}.$$

With the mgf, we can readily obtain $\mathrm{E}(Y_1|Y_1 \in \mathcal{A}) = \mu_1 - \sigma_1\mathsf{P}$ and $\mathrm{Var}(Y_1|Y_1 \in \mathcal{A}) = \sigma_1^2(1 - \mathsf{Q} - \mathsf{P}^2)$ (Johnson et al., 1994, p. 156–158), where

$$\mathsf{P} = \frac{\phi(\tau_b) - \phi(\tau_a)}{\Phi(\tau_b) - \Phi(\tau_a)}; \quad \mathsf{Q} = \frac{\tau_b\phi(\tau_b) - \tau_a\phi(\tau_a)}{\Phi(\tau_b) - \Phi(\tau_a)}.$$

Now, we define $Y_2 = \mu_2 + \sigma_2 X$, and then Y_2 follows a normal distribution with mean μ_2 and variance σ_2^2, i.e., $Y_2 \sim N(\mu_2, \sigma_2^2)$. Assume that Y_1 and Y_2 jointly follow a bivariate normal distribution with correlation ρ, represented by $(Y_1, Y_2) \sim f(y_1, y_2; \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. We are interested in the distribution of Y_2 and relevant statistic, such as correlation between Y_2 and Y_1, given that Y_1 falls within the interval $\mathcal{A}$. Here, truncation is applied only to Y_1, while Y_2 remains untruncated, that is, $a \le Y_1 \le b$ and $-\infty < Y_2 < \infty$.

In the following, we presented two methods for deriving these expected results, both of which use the findings in Kotz et al. (2000, p. 311–312, 315) as a starting point.

A.1 Method 1: based on the technique from Kotz et al. (2000)

This is a relatively straightforward technique without intensive computation. For the bivariate normal distribution of (Y_1, Y_2), the conditional distribution of Y_2 given $Y_1 = y_1$ is also normally distributed, and can be expressed as follows:

$$Y_2|Y_1 = y_1 \sim N\left(\mu_2 - \rho\mu_1\frac{\sigma_2}{\sigma_1} + \rho\frac{\sigma_2}{\sigma_1}y_1, \sigma_2^2(1-\rho^2)\right).$$

We then have the following identities

$$\begin{aligned}\mathrm{E}(Y_2) &= \mathrm{E}[\mathrm{E}(Y_2|Y_1)] = \mathrm{E}\left(\mu_2 - \rho\mu_1\frac{\sigma_2}{\sigma_1} + \rho\frac{\sigma_2}{\sigma_1}Y_1\right)\\ &= \mu_2 - \rho\mu_1\frac{\sigma_2}{\sigma_1} + \rho\frac{\sigma_2}{\sigma_1}\mathrm{E}(Y_1),\\ \mathrm{E}(Y_1Y_2) &= \mathrm{E}(Y_1\mathrm{E}(Y_2|Y_1)) = \mathrm{E}\left[Y_1\left(\mu_2 - \rho\mu_1\frac{\sigma_2}{\sigma_1} + \rho\frac{\sigma_2}{\sigma_1}Y_1\right)\right]\\ &= \left(\mu_2 - \rho\mu_1\frac{\sigma_2}{\sigma_1}\right)\mathrm{E}(Y_1) + \rho\frac{\sigma_2}{\sigma_1}\mathrm{E}(Y_1^2),\end{aligned}$$

$$
\begin{aligned}
\mathrm{E}(Y_2^2) &= \mathrm{E}(\mathrm{E}(Y_2^2|Y_1)) = \mathrm{E}\left[\mathrm{Var}(Y_2|Y_1) + (\mathrm{E}(Y_2|Y_1))^2\right] \\
&= \sigma_2^2(1-\rho^2) + \left(\mu_2 - \rho\mu_1\frac{\sigma_2}{\sigma_1}\right)^2 \\
&\quad + 2\left(\mu_2 - \rho\mu_1\frac{\sigma_2}{\sigma_1}\right)\rho\frac{\sigma_2}{\sigma_1}\mathrm{E}(Y_1) + \rho^2\frac{\sigma_2^2}{\sigma_1^2}\mathrm{E}(Y_1^2), \\
\mathrm{Var}(Y_2) &= \mathrm{E}(Y_2^2) - (\mathrm{E}(Y_2))^2 = \sigma_2^2(1-\rho^2) + \rho^2\frac{\sigma_2^2}{\sigma_1^2}\left[\mathrm{E}(Y_1^2) - (\mathrm{E}(Y_1))^2\right] \\
&= \sigma_2^2\left[(1-\rho^2) + \frac{\rho^2}{\sigma_1^2}\mathrm{Var}(Y_1)\right], \\
\mathrm{Cov}(Y_1, Y_2) &= \mathrm{E}(Y_1Y_2) - \mathrm{E}(Y_1)\mathrm{E}(Y_2) = \rho\frac{\sigma_2}{\sigma_1}\left[\mathrm{E}(Y_1^2) - (\mathrm{E}(Y_1))^2\right] = \rho\frac{\sigma_2}{\sigma_1}\mathrm{Var}(Y_1), \\
\mathrm{Corr}(Y_1, Y_2) &= \frac{\mathrm{Cov}(Y_1, Y_2)}{\sqrt{\mathrm{Var}(Y_1)}\sqrt{\mathrm{Var}(Y_2)}} = \rho\sqrt{\left(\rho^2 + \frac{(1-\rho^2)\sigma_1^2}{\mathrm{Var}(Y_1)}\right)^{-1}}.
\end{aligned}
$$

Clearly, by substituting $\mathrm{E}(Y_1)$ and $\mathrm{Var}(Y_1)$ in these identities with the previously derived expressions $\mathrm{E}(Y_1|Y_1 \in \mathcal{A}) = \mu_1 - \sigma_1\mathsf{P}$ and $\mathrm{Var}(Y_1|Y_1 \in \mathcal{A}) = \sigma_1^2(1-\mathsf{Q}-\mathsf{P}^2)$ from the univariate truncated distribution, we can readily obtain the desired results for the distribution of Y_2 given that Y_1 falls within the interval $\mathcal{A}$. These include $\mathrm{E}(Y_2|Y_1 \in \mathcal{A})$, $\mathrm{Var}(Y_2|Y_1 \in \mathcal{A})$, and $\mathrm{Corr}(Y_1, Y_2|Y_1 \in \mathcal{A})$. These results are identical to those obtained using an alternative method involving the mgf, which will be presented subsequently.

A.2 Method 2: based on moment generating function

With the specified truncation on Y_1 as $a \le Y_1 \le b$, the pdf of the resulting truncated distribution for (Y_1, Y_2) is then expressed as $f(y_1, y_2; \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho | a \le y_1 \le b) = f(y_1, y_2)[\Phi(\tau_b) - \Phi(\tau_a)]^{-1}$, and the mgf of this truncated bivariate distribution for (Y_1, Y_2) can then be derived as follows:

$$
\begin{aligned}
M_{(Y_1,Y_2)}(t_1, t_2) &= \mathrm{E}\left(\mathrm{e}^{t_1Y_1 + t_2Y_2}|Y_1 \in A\right) \\
&= \left[\Phi(\tau_b) - \Phi(\tau_a)\right]^{-1}\int_{-\infty}^{\infty}\int_a^b \mathrm{e}^{t_1y_1 + t_2y_2}f(y_1, y_2)\mathrm{d}y_1\mathrm{d}y_2 \\
&= \widetilde{\mathsf{R}}\exp(\mathsf{A}),
\end{aligned}
$$

where

$$
\begin{aligned}
\mathsf{A} &= t_1\mu_1 + t_2\mu_2 + \frac{1}{2}\left(t_1^2\sigma_1^2 + 2t_1t_2\rho\sigma_1\sigma_2 + t_2^2\sigma_2^2\right), \\
\widetilde{\mathsf{R}} &= \left[\Phi\left(\frac{b-\widetilde{\mu}_1}{\sigma_1}\right) - \Phi\left(\frac{a-\widetilde{\mu}_1}{\sigma_1}\right)\right]\left[\Phi(\tau_b) - \Phi(\tau_a)\right]^{-1}, \\
\widetilde{\mu}_1 &= \mu_1 + t_1\sigma_1^2 + t_2\rho\sigma_1\sigma_2.
\end{aligned}
$$

From the mgf, the first, the second, and mixed derivatives in terms of t_1 and t_2 can be obtained as follows:

$$
\begin{aligned}
\frac{\partial}{\partial t_1}M_{(Y_1,Y_2)}(t_1, t_2) &= \widetilde{\mu}_1\widetilde{\mathsf{R}}\exp(\mathsf{A}) - \sigma_1\widetilde{\mathsf{P}}\exp(\mathsf{A}), \\
\frac{\partial}{\partial t_2}M_{(Y_1,Y_2)}(t_1, t_2) &= \widetilde{\mu}_2\widetilde{\mathsf{R}}\exp(\mathsf{A}) - \rho\sigma_2\widetilde{\mathsf{P}}\exp(\mathsf{A}), \\
\frac{\partial^2}{\partial t_1^2}M_{(Y_1,Y_2)}(t_1, t_2) &= \sigma_1^2\widetilde{\mathsf{R}}\exp(\mathsf{A}) + \widetilde{\mu}_1\left[\widetilde{\mu}_1\widetilde{\mathsf{R}}\exp(\mathsf{A}) - \sigma_1\widetilde{\mathsf{P}}\exp(\mathsf{A})\right] \\
&\quad - \left[\widetilde{\mu}_1\sigma_1\widetilde{\mathsf{P}}\exp(\mathsf{A}) + \sigma_1^2\widetilde{\mathsf{Q}}\exp(\mathsf{A})\right], \\
\frac{\partial^2}{\partial t_2^2}M_{(Y_1,Y_2)}(t_1, t_2) &= \sigma_2^2\widetilde{\mathsf{R}}\exp(\mathsf{A}) + \widetilde{\mu}_2\left[\widetilde{\mu}_2\widetilde{\mathsf{R}}\exp(\mathsf{A}) - \rho\sigma_2\widetilde{\mathsf{P}}\exp(\mathsf{A})\right] \\
&\quad - \left[\widetilde{\mu}_2\rho\sigma_2\widetilde{\mathsf{P}}\exp(\mathsf{A}) + \rho^2\sigma_2^2\widetilde{\mathsf{Q}}\exp(\mathsf{A})\right], \\
\frac{\partial^2}{\partial t_1\partial t_2}M_{(Y_1,Y_2)}(t_1, t_2) &= \rho\sigma_1\sigma_2\widetilde{\mathsf{R}}\exp(\mathsf{A}) + \widetilde{\mu}_1\left[\widetilde{\mu}_2\widetilde{\mathsf{R}}\exp(\mathsf{A}) - \rho\sigma_2\widetilde{\mathsf{P}}\exp(\mathsf{A})\right] \\
&\quad - \left[\widetilde{\mu}_2\sigma_1\widetilde{\mathsf{P}}\exp(\mathsf{A}) + \rho\sigma_1\sigma_2\widetilde{\mathsf{Q}}\exp(\mathsf{A})\right],
\end{aligned}
$$

where

$$\widetilde{\mu}_2 = \mu_2 + t_2\sigma_2^2 + t_1\rho\sigma_1\sigma_2,$$
$$\widetilde{\mathsf{P}} = \left[\phi\left(\frac{b-\widetilde{\mu}_1}{\sigma_1}\right) - \phi\left(\frac{a-\widetilde{\mu}_1}{\sigma_1}\right)\right]\left[\Phi\left(\tau_b\right) - \Phi\left(\tau_a\right)\right]^{-1},$$
$$\widetilde{\mathsf{Q}} = \left[\left(\frac{b-\widetilde{\mu}_1}{\sigma_1}\right)\phi\left(\frac{b-\widetilde{\mu}_1}{\sigma_1}\right) - \left(\frac{a-\widetilde{\mu}_1}{\sigma_1}\right)\phi\left(\frac{a-\widetilde{\mu}_1}{\sigma_1}\right)\right]\left[\Phi\left(\tau_b\right) - \Phi\left(\tau_a\right)\right]^{-1}.$$

By substituting $t_1 = 0$ and $t_2 = 0$ into these identities, we have $\mathsf{A} = 0$ and $\widetilde{\mu}_1 = \mu_1$. This allows us to easily derive the desired results as below for the distribution of both Y_2 and Y_1, given that Y_1 falls within the interval $\mathcal{A}$.

$$\mathrm{E}(Y_1|Y_1 \in \mathcal{A}) = \frac{\partial}{\partial t_1}M_{(Y_1,Y_2)}(t_1,t_2)\Big|_{t_1=0,t_2=0} = \mu_1 - \sigma_1\mathsf{P},$$
$$\mathrm{E}(Y_2|Y_1 \in \mathcal{A}) = \frac{\partial}{\partial t_2}M_{(Y_1,Y_2)}(t_1,t_2)\Big|_{t_1=0,t_2=0} = \mu_2 - \rho\sigma_2\mathsf{P},$$
$$\mathrm{E}(Y_1^2|Y_1 \in \mathcal{A}) = \frac{\partial^2}{\partial t_1^2}M_{(Y_1,Y_2)}(t_1,t_2)\Big|_{t_1=0,t_2=0} = \sigma_1^2 + \mu_1^2 - 2\mu_1\sigma_1\mathsf{P} - \sigma_1^2\mathsf{Q},$$
$$\mathrm{E}(Y_2^2|Y_1 \in \mathcal{A}) = \frac{\partial^2}{\partial t_2^2}M_{(Y_1,Y_2)}(t_1,t_2)\Big|_{t_1=0,t_2=0} = \sigma_2^2 + \mu_2^2 - 2\rho\sigma_2\mu_2\mathsf{P} - \rho^2\sigma_2^2\mathsf{Q},$$
$$\begin{aligned}\mathrm{E}(Y_1Y_2|Y_1 \in \mathcal{A}) &= \frac{\partial^2}{\partial t_1\partial t_2}M_{(Y_1,Y_2)}(t_1,t_2)\Big|_{t_1=0,t_2=0}\\ &= \mu_2\left(\mu_1 - \sigma_1\mathsf{P}\right) + \rho\sigma_1\sigma_2\left(1 - \frac{\mu_1}{\sigma_1}\mathsf{P} - \mathsf{Q}\right),\end{aligned}$$
$$\mathrm{Var}(Y_1|Y_1 \in \mathcal{A}) = \mathrm{E}(Y_1^2|Y_1 \in \mathcal{A}) - [\mathrm{E}(Y_1|Y_1 \in \mathcal{A})]^2 = \sigma_1^2\left(1 - \mathsf{Q} - \mathsf{P}^2\right),$$
$$\mathrm{Var}(Y_2|Y_1 \in \mathcal{A}) = \mathrm{E}(Y_2^2|Y_1 \in \mathcal{A}) - [\mathrm{E}(Y_2|Y_1 \in \mathcal{A})]^2 = \sigma_2^2\left(1 - \rho^2\mathsf{Q} - \rho^2\mathsf{P}^2\right),$$
$$\begin{aligned}\mathrm{Cov}(Y_1, Y_2|Y_1 \in \mathcal{A}) &= \mathrm{E}(Y_1Y_2|Y_1 \in \mathcal{A}) - \mathrm{E}(Y_1|Y_1 \in \mathcal{A})\mathrm{E}(Y_2|Y_1 \in \mathcal{A})\\ &= \rho\sigma_1\sigma_2\left(1 - \mathsf{Q} - \mathsf{P}^2\right),\end{aligned}$$
$$\begin{aligned}\mathrm{Corr}(Y_1, Y_2|Y_1 \in \mathcal{A}) &= \frac{\mathrm{Cov}(Y_1, Y_2|Y_1 \in \mathcal{A})}{\sqrt{\mathrm{Var}(Y_1|Y_1 \in \mathcal{A})}\sqrt{\mathrm{Var}(Y_2|Y_1 \in \mathcal{A})}}\\ &= \rho\sqrt{\left(\rho^2 + \frac{1-\rho^2}{1-\mathsf{Q}-\mathsf{P}^2}\right)^{-1}}.\end{aligned}$$

A.3 Single Truncation

Up to this point, our discussion has focussed on the double trunction of Y_1 wtihin the interval $\mathcal{A} = [a, b]$. Now, let's examine its special cases of single truncation for Y_1. There are two specific scenarios: (1) setting $a = -\infty$, which results in right truncation; (2) letting $b = \infty$, which leads to left truncation. The previously discussed results regarding the truncated distribution for both Y_1 and Y_2 apply to all forms of truncation of Y_1, as long as Y_2 remains untruncated. The key difference is that, for right trunction, P and Q should be replaced with P_r and Q_r, respectively, as defined below. Similarly, for left truncation, P and Q should be substituted with P_l and Q_l, respectively.

$$\text{Right truncation, } \mathcal{A} = (-\infty, b]: \mathsf{P}_r = \frac{\phi\left(\tau_b\right)}{\Phi\left(\tau_b\right)}; \quad \mathsf{Q}_r = \frac{\tau_b\phi\left(\tau_b\right)}{\Phi\left(\tau_b\right)};$$
$$\text{Left truncation, } \mathcal{A} = [a, \infty): \mathsf{P}_l = -\frac{\phi\left(\tau_a\right)}{1-\Phi\left(\tau_a\right)}; \quad \mathsf{Q}_l = -\frac{\tau_a\phi\left(\tau_a\right)}{1-\Phi\left(\tau_a\right)}.$$