Let X^+, $X^+ \sim N(\mu_{X^+}, \sigma^2)$ be response in subgroup A^+, and X^-, $X^- \sim N(\mu_{X^-}, \sigma^2)$ be response in subgroup A^-. Let n_1 and n_2 be the sample sizes in A- and A+ subgroups. Denote $\bar{X}^+$ to be the sample average of responses in the biomarker positive subgroup and $\bar{X}^-$ be the sample average of responses in the biomarker negative subgroup.

$$\begin{pmatrix} \bar{X}^+ \\ \bar{X}^- \end{pmatrix} \sim \left(N\begin{pmatrix} \mu_{X^+} \\ \mu_{X^-} \end{pmatrix}, \begin{pmatrix} \dfrac{\sigma^2}{n_1} & 0 \\ 0 & \dfrac{\sigma^2}{n_2} \end{pmatrix} \right)$$

We apply isotonic regression when $\overline{X}^- > \overline{X}^+$. Let $p = n_1 / (n_1 + n_2)$ be the sampling proportion of A^+ in the cohort, Then the sample average in the pooled sample of A^+ and A^- is $X = p\bar{X}^+ + (1-p)\bar{X}^-$. We will keep the raw estimate if we observe $\overline{X}^+ \geq \overline{X}^-$.

The goal is to find the distribution of $\bar{X}^+$ and $\bar{X}^-$ when $\overline{X}^+ \geq \overline{X}^-$ $X^+ \geq X^-$, and the distribution of $\bar{X}$ when $\overline{X}^- > \overline{X}^+$ $X^- > X^+$.

Now, let $W = \overline{X}^+ - \overline{X}^-$, the conditional distribution of $X^+ | W \geq 0$ can be derived from:

$$\begin{pmatrix} \overline{X}^+ \\ W \end{pmatrix} \sim \left(N\begin{pmatrix} \mu_{X^+} \\ \mu_{X^+} - \mu_{X^-} \end{pmatrix}, \begin{pmatrix} \dfrac{\sigma^2}{n_1} & \dfrac{\sigma^2}{n_1} \\ \dfrac{\sigma^2}{n_1} & \dfrac{\sigma^2}{n_1} + \dfrac{\sigma^2}{n_2} \end{pmatrix} \right) \tag{1}$$

Using the formula in the appendix B of the paper, we can calculate from (1) that:

$$
\begin{aligned}
E[\overline{X}^+ \,|\, W \geq 0] &= E\left[E[\overline{X}^+ \,|\, W \geq 0, W] \right] \\
&= E\left[E[\overline{X}^+ \,|\, W] \,|\, W > 0 \right] \\
&= E\left[\mu_{X^+} + \frac{\sqrt{\frac{\sigma^2}{n_1}}}{\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}} \frac{\frac{\sigma^2}{n_1}}{\sqrt{\frac{\sigma^2}{n_1}}\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}} (W-(\mu_{X^+}-\mu_{X^-})) \Big| W > 0 \right] \\
&= E\left[\mu_{X^+} + \frac{\frac{\sigma^2}{n_1}}{\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}} \frac{W-(\mu_{X^+}-\mu_{X^-})}{\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}} \middle| \frac{W-(\mu_{X^+}-\mu_{X^-})}{\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}} > \frac{-(\mu_{X^+}-\mu_{X^-})}{\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}} \right] \\
&= \mu_{X^+} + \frac{\frac{\sigma^2}{n_1}}{\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}} E\left[Z \Big| Z > \frac{-(\mu_{X^+}-\mu_{X^-})}{\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}} \right] \\
&= \mu_{X^+} + \sigma\sqrt{\frac{n_2}{(n_1+n_2)n_1}} E\left[Z \Big| Z > \frac{-(\mu_{X^+}-\mu_{X^-})}{\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}} \right]
\end{aligned}
$$

where Z is a standard normal random variable and define

$$
h_m(t) = E[Z^m \,|\, Z \geq t] = \{1-\Phi(t)\}^{-1} \int_t^{+\infty} u^m \phi(u)\,du
$$

m = 1,2, where $\phi(.)$ is the density function of a standard normal random variable and $\Phi(.)$ is the corresponding cumulative distribution function. Then

$$
E[\overline{X}^+ \,|\, W \geq 0] = \mu_{X^+} + \sigma\sqrt{\frac{n_2}{(n_1+n_2)n_1}} h_1(w^*),
$$

where $w^* = \dfrac{-(\mu_{X^+} - \mu_{X^-})}{\sqrt{\dfrac{\sigma^2}{n_1} + \dfrac{\sigma^2}{n_2}}}$, and the conditional variance is

$$Var[\overline{X}^+ \mid W \geq 0] = \frac{\sigma^2}{n_1}\left[1 - \frac{n_2}{n_1 + n_2}\left\{1 - h_2(w^*) + \left(h_1(w^*)\right)^2\right\}\right].$$

Similarly, the joint distribution of $X^- \mid W \geq 0$ is derived from the joint distribution

$$\begin{pmatrix} \overline{X}^- \\ W \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_{X^-} \\ \mu_{X^+} - \mu_{X^-} \end{pmatrix}, \begin{pmatrix} \dfrac{\sigma^2}{n_2} & -\dfrac{\sigma^2}{n_2} \\ -\dfrac{\sigma^2}{n_2} & \dfrac{\sigma^2}{n_1} + \dfrac{\sigma^2}{n_2} \end{pmatrix}\right) \qquad (2)$$

Using similar method outlined above, we can get:

$$E[\overline{X}^- \mid W \geq 0] = \mu_{X^-} - \sigma\sqrt{\frac{n_1}{(n_1 + n_2)n_2}}h_1(w^*),$$

where $w^* = \dfrac{-(\mu_{X^+} - \mu_{X^-})}{\sqrt{\dfrac{\sigma^2}{n_1} + \dfrac{\sigma^2}{n_2}}}$, and the conditional variance is

$$Var[\overline{X}^- \mid W \geq 0] = \frac{\sigma^2}{n_2}\left[1 - \frac{n_1}{n_1 + n_2}\left\{1 - h_2(w^*) + \left(h_1(w^*)\right)^2\right\}\right].$$

The joint distribution related to $X \mid W < 0$ is

$$\begin{pmatrix} X \\ W \end{pmatrix} \sim N\left(\begin{pmatrix} \dfrac{n_1}{n_1 + n_2}\mu_{X^+} + \dfrac{n_2}{n_1 + n_2}\mu_{X^-} \\ \mu_{X^+} - \mu_{X^-} \end{pmatrix}, \begin{pmatrix} \dfrac{\sigma^2}{n_1 + n_2} & 0 \\ 0 & \dfrac{\sigma^2}{n_1} + \dfrac{\sigma^2}{n_2} \end{pmatrix}\right) \qquad (3)$$

We find that they are independent random variable. Thus, we can write the conditional mean and variance as

$$E[X \mid W < 0] = \frac{n_1}{n_1 + n_2}\mu_{X^+} + \frac{n_2}{n_1 + n_2}\mu_{X^-},$$

$$Var[X \mid W < 0] = \frac{\sigma^2}{n_1 + n_2}$$

Note that, from Johnson and Kotz (1970, pp 81-83), $h_1(t) = \phi(t)/\{1 - \Phi(t)\}$ and $h_2(t) = 1 + t \times h_1(t)$. Under the null hypothesis, $\mu_{X^+} = \mu_{X^-}$, thus $w = 0$, $h_1(0) = 2\phi(0)$ and $h_2(0) = 1$. Thus, do not pull the response rate,

$$E[\overline{X}^+ \mid W \geq 0] = \mu_{X^+} + 2\phi(0)\sigma\sqrt{\frac{n_2}{(n_1 + n_2)n_1}}$$

$$Var[\overline{X}^+ \mid W \geq 0] = \frac{\sigma^2}{n_1}\left[1 - 4\phi^2(0)\frac{n_2}{n_1 + n_2}\right]$$

$$E[\overline{X}^- \mid W \geq 0] = \mu_{X^-} - 2\phi(0)\sigma\sqrt{\frac{n_1}{(n_1 + n_2)n_2}}$$

$$Var[\overline{X}^+ \mid W \geq 0] = \frac{\sigma^2}{n_1}\left[1 - 4\phi^2(0)\frac{n_1}{n_1 + n_2}\right]$$