5.2.5.5 の最初にある一般的な混合分布について確認してみる。
一般的に混合分布は以下のように表される。

$$p(x_{mix}) = \sum_k p(x_{mix}|k)p(k) = \sum_k \pi_k p(x_{mix}|k) \tag{1}$$

このとき、平均を求めると

$$\mathbb{E}[x_{mix}] = \int x_{mix} p(x_{mix}) dx_{mix} = \int x_{mix} \sum_k \pi_k p(x_{mix}|k) dx_{mix} = \sum_k \pi_k \int x_{mix} p(x_{mix}|k) dx_{mix} = \sum_k \pi_k \mu_k \tag{2}$$

同様に分散を求めると

$$\begin{aligned}
\mathbb{V}_{p(x_{mix})}[x_{mix}] = \mathbb{E}_{p(x_{mix})}[x_{mix}^2] - \mathbb{E}[x_{mix}]^2 = \int x_{mix}^2 p(x_{mix}) dx_{mix} - \mathbb{E}[x_{mix}]^2 \\
= \int x_{mix}^2 \sum_k \pi_k p(x_{mix}|k) dx_{mix} - \mathbb{E}[x_{mix}]^2 = \sum_k (\pi_k \int x_{mix}^2 p(x_{mix}|k) dx_{mix}) - \mathbb{E}[x_{mix}]^2 \\
= \sum_k (\pi_k \mathbb{E}_{p(x_{mix}|k)}[x_{mix}^2]) - \mathbb{E}[x_{mix}]^2 = \sum_k (\pi_k \mathbb{E}_{p(x_{mix}|k)}[x_{mix}^2]) - \mathbb{E}[x_{mix}]^2 \\
= \sum_k (\pi_k (\mathbb{E}_{p(x_{mix}|k)}[(x_{mix} - \mathbb{E}_{p(x_{mix}|k)}[x_{mix}])^2] + \mathbb{E}_{p(x_{mix}|k)}[x_{mix}]^2)) - \mathbb{E}[x_{mix}]^2 \\
= \sum_k (\pi_k (v_k + {\mu_k}^2)) - \mathbb{E}[x_{mix}]^2
\end{aligned} \tag{3}$$

上記になるが、本の記載は少し異なっているように思う。
活性の分布を考えてみる。
今回、活性化関数として ReLU を選択している。もともと、正規分布だったものを、0 以下を 0 にして、それ以外をそのままとするので、入力が正規分布だとすると、本にあるように ReLU を通ったあとの分布は、平均 0 と分散 0 の質点と 0 以上の部分の混合分布となる。
混合比はもとのガウス分布の 0 以下の部分の面積と 0 以上の部分の面積になる。0 以下の面積は (5.70) にかかれている内容になり、0 以上の部分は残りなので、(5.71) になる。
一般に平均と分散が $\mu, \sigma^2$ で切断位置が a の場合、それぞれの混合係数は、

$$\pi_{low} = \int_{-\infty}^{a} \mathcal{N}(x|\mu, \sigma^2) dx = \int_{-\infty}^{\frac{a-\mu}{\sigma}} \mathcal{N}(y|0,1) dy = \Phi(\frac{a-\mu}{\sigma}) \tag{4}$$

ただし、$y = \frac{x-\mu}{\sigma}$ で変数変換をしている。$\Phi$ は (2.26) に定義がある。
残りの部分は、正規分布の形状を考えると、

$$\pi_{high} = 1 - \Phi(\frac{a-\mu}{\sigma}) = \Phi(-\frac{a-\mu}{\sigma}) \tag{5}$$

切断ガウス分布の平均を愚直に求める。(変数はスカラ) そもそも、切断ガウス分布は以下のようになる。

$$p(x|x > a) = \frac{1}{Z} \mathcal{N}(x|\mu, \sigma) \tag{6}$$

基本的にガウス分布だが、一部しか利用していないため、全体を積分すると 1 にならないため、正規化定数 Z が必要。

$$Z = \int_{a}^{\infty} \mathcal{N}(x|\mu, \sigma) dx = 1 - \Phi(\frac{a-\mu}{\sigma}) \tag{7}$$

これを踏まえて、平均値を計算する。

$$\begin{aligned}Z\mathbb{E}[x|x>a] = \int_a^\infty x\mathcal{N}(x|\mu,\sigma)dx = \frac{1}{\sqrt{2\pi\sigma^2}}\int_a^\infty xexp(-\frac{(x-\mu)^2}{2\sigma^2})dx = \frac{1}{\sqrt{2\pi\sigma^2}}\int_a^\infty (x-\mu+\mu)exp(-\frac{(x-\mu)^2}{2\sigma^2})dx \\ = \frac{1}{\sqrt{2\pi\sigma^2}}(\int_a^\infty (x-\mu)exp(-\frac{(x-\mu)^2}{2\sigma^2})dx + \int_a^\infty \mu exp(-\frac{(x-\mu)^2}{2\sigma^2})dx) \\ = \frac{1}{\sqrt{2\pi}}(\int_a^\infty \frac{x-\mu}{\sigma}exp(-\frac{(x-\mu)^2}{2\sigma^2})\sigma\frac{dx}{\sigma} + \int_a^\infty \mu exp(-\frac{(x-\mu)^2}{2\sigma^2})\frac{dx}{\sigma})\end{aligned} \tag{8}$$

ここで、$y=\frac{x-\mu}{\sigma}$ とすると、(5.48) の重積分の変数変換も参考にして、

$$\begin{aligned}Z\mathbb{E}[x|x>a] = \frac{1}{\sqrt{2\pi}}\left(\sigma\int_{\frac{a-\mu}{\sigma}}^\infty yexp\left(-\frac{y^2}{2}\right)dy + \mu\int_{\frac{a-\mu}{\sigma}}^\infty exp\left(-\frac{y^2}{2}\right)dy\right) \\ = \frac{1}{\sqrt{2\pi}}\left(\sigma\int_{\frac{a-\mu}{\sigma}}^\infty yexp\left(-\frac{y^2}{2}\right)dy + \mu\int_{\frac{a-\mu}{\sigma}}^\infty exp\left(-\frac{y^2}{2}\right)dy\right) \\ = \sigma\left[\frac{1}{\sqrt{2\pi}}exp\left(-\frac{y^2}{2}\right)\right]_\infty^{\frac{a-\mu}{\sigma}} + \mu(1-\Phi(\frac{a-\mu}{\sigma})) = \sigma\mathcal{N}\left(\frac{a-\mu}{\sigma}|0,1\right) + \mu(1-\Phi(\frac{a-\mu}{\sigma}))\end{aligned} \tag{9}$$

両辺を Z で割ると、

$$\mu_t = \mathbb{E}[x|x>a] = \sigma\frac{\mathcal{N}\left(\frac{a-\mu}{\sigma}|0,1\right)}{(1-\Phi(\frac{a-\mu}{\sigma}))} + \mu = \mu + \sigma\frac{\mathcal{N}\left(\frac{a-\mu}{\sigma}|0,1\right)}{\Phi(-\frac{a-\mu}{\sigma})} \tag{10}$$

(5.72) が求まった。

同様に分散を求める。まず 2 次モーメントを求める。

$$\begin{aligned}Z\mathbb{E}[x^2|x>a] = \int_a^\infty x^2\mathcal{N}(x|\mu,\sigma)dx = \frac{1}{\sqrt{2\pi\sigma^2}}\int_a^\infty x^2exp(-\frac{(x-\mu)^2}{2\sigma^2})dx \\ = \frac{1}{\sqrt{2\pi\sigma^2}}\int_a^\infty ((x-\mu)^2+2\mu x-\mu^2)exp(-\frac{(x-\mu)^2}{2\sigma^2})dx \\ = \frac{1}{\sqrt{2\pi\sigma^2}}(\int_a^\infty (x-\mu)^2exp(-\frac{(x-\mu)^2}{2\sigma^2})dx + \int_a^\infty 2\mu xexp(-\frac{(x-\mu)^2}{2\sigma^2})dx - \int_a^\infty \mu^2exp(-\frac{(x-\mu)^2}{2\sigma^2})dx)\end{aligned} \tag{11}$$

ここで、$y=\frac{x-\mu}{\sigma}$ とすると、(5.48) の重積分の変数変換も参考にして、

$$\begin{aligned}Z\mathbb{E}[x^2|x>a] = \int_{\frac{a-\mu}{\sigma}}^\infty \frac{1}{\sqrt{2\pi}}\sigma^2y^2exp(-\frac{y^2}{2})dy + 2\mu(\sigma\mathcal{N}\left(\frac{a-\mu}{\sigma}|0,1\right) + \mu(1-\Phi(\frac{a-\mu}{\sigma}))) - \mu^2\int_{\frac{a-\mu}{\sigma}}^\infty \frac{1}{\sqrt{2\pi}}exp(-\frac{y^2}{2})dy \\ = \frac{1}{\sqrt{2\pi}}\int_{\frac{a-\mu}{\sigma}}^\infty \sigma^2y^2exp(-\frac{y^2}{2})dy + 2\mu(\sigma\mathcal{N}\left(\frac{a-\mu}{\sigma}|0,1\right) + \mu(1-\Phi(\frac{a-\mu}{\sigma}))) - \mu^2(1-\Phi(\frac{a-\mu}{\sigma})) \\ = \frac{1}{\sqrt{2\pi}}\int_{\frac{a-\mu}{\sigma}}^\infty \sigma^2y^2exp(-\frac{y^2}{2})dy + 2\mu\sigma\mathcal{N}\left(\frac{a-\mu}{\sigma}|0,1\right) + \mu^2(1-\Phi(\frac{a-\mu}{\sigma})) \\ = \frac{\sigma^2}{\sqrt{2\pi}}\left(\left[-yexp(-\frac{y^2}{2})\right]_{\frac{a-\mu}{\sigma}}^\infty + \int_{\frac{a-\mu}{\sigma}}^\infty exp(-\frac{y^2}{2})dy\right) + 2\mu\sigma\mathcal{N}\left(\frac{a-\mu}{\sigma}|0,1\right) + \mu^2(1-\Phi(\frac{a-\mu}{\sigma})) \\ = \sigma^2\left(\frac{a-\mu}{\sigma}\mathcal{N}(\frac{a-\mu}{\sigma}|0,1) + (1-\Phi(\frac{a-\mu}{\sigma}))\right) + 2\mu\sigma\mathcal{N}\left(\frac{a-\mu}{\sigma}|0,1\right) + \mu^2(1-\Phi(\frac{a-\mu}{\sigma}))\end{aligned} \tag{12}$$

よって、

$$
\begin{aligned}
&\mathbb{E}[(x-\mu_t)^2|x>a] = \mathbb{E}[x^2|x>a] - (\mathbb{E}[x^2|x>a])^2 \\
&= \frac{1}{1-\Phi(\frac{a-\mu}{\sigma})}\left(\sigma^2\left(\frac{a-\mu}{\sigma}\mathcal{N}(\frac{a-\mu}{\sigma}|0,1)+(1-\Phi(\frac{a-\mu}{\sigma}))\right)+2\mu\sigma\mathcal{N}\left(\frac{a-\mu}{\sigma}|0,1\right)+\mu^2(1-\Phi(\frac{a-\mu}{\sigma}))\right) \\
&\quad -(\mu+\sigma\frac{\mathcal{N}\left(\frac{a-\mu}{\sigma}|0,1\right)}{1-\Phi(\frac{a-\mu}{\sigma})})^2 \\
&= \left(\sigma^2\left(\frac{1}{1-\Phi(\frac{a-\mu}{\sigma})}\frac{a-\mu}{\sigma}\mathcal{N}(\frac{a-\mu}{\sigma}|0,1)+1\right)+\frac{2\mu\sigma}{1-\Phi(\frac{a-\mu}{\sigma})}\mathcal{N}\left(\frac{a-\mu}{\sigma}|0,1\right)+\mu^2\right) \\
&\quad -(\mu^2+2\mu\sigma\frac{\mathcal{N}\left(\frac{a-\mu}{\sigma}|0,1\right)}{1-\Phi(\frac{a-\mu}{\sigma})}+(\sigma\frac{\mathcal{N}\left(\frac{a-\mu}{\sigma}|0,1\right)}{1-\Phi(\frac{a-\mu}{\sigma})})^2) \\
&= \sigma^2\left(\frac{1}{1-\Phi(\frac{a-\mu}{\sigma})}\frac{a-\mu}{\sigma}\mathcal{N}(\frac{a-\mu}{\sigma}|0,1)+1\right)-(\sigma\frac{\mathcal{N}\left(\frac{a-\mu}{\sigma}|0,1\right)}{1-\Phi(\frac{a-\mu}{\sigma})})^2 = \sigma^2\left(1+\overline{\mu}\frac{\mathcal{N}(\overline{\mu}|0,1)}{\Phi(-\overline{\mu})}-\left(\frac{\mathcal{N}\left(\overline{\mu}|0,1\right)}{\Phi(-\overline{\mu}})\right)^2\right)
\end{aligned}
\tag{13}
$$

となる。これは本とは少し異なるが、wikipedia の記載 (https://ja.wikipedia.org/wiki/切断正規分布) には一致している。