# Background

Before we prove some key results for truncated Gausians, we recall the definition and key properties of the one-dimensional Gaussian distribution.

**Definition 1** (One-Dimensional Gaussian Distribution). Given parameters $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$, $\tau \in \mathbb{R}$, and $\rho \in \mathbb{R}^+$, a one-dimensional random variable $X \in \mathbb{R}$ is distributed according to a one-dimensional Gaussian distribution if the density has the form $\mathcal{N}\left(\cdot;\mu,\sigma^2\right)$ or $\mathcal{G}\left(\cdot;\tau,\rho\right)$ with

$$\mathcal{N}\left(x;\mu,\sigma^2\right) := \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \tag{1}$$

$$\mathcal{G}\left(x;\tau,\rho\right) := \sqrt{\frac{\rho}{2\pi}}\cdot\exp\left(-\frac{\tau^2}{2\rho}\right)\cdot\exp\left(\tau\cdot x-\rho\cdot\frac{x^2}{2}\right). \tag{2}$$

If $\mu = \tau = 0$ and $\sigma^2 = \rho = 1$ we simply write

$$\mathcal{N}\left(x\right) := \mathcal{N}\left(x;0,1\right), \tag{3}$$

$$\mathcal{G}\left(x\right) := \mathcal{G}\left(x;0,1\right). \tag{4}$$

Note that both definitions are identical when using the identities

$$\tau = \frac{\mu}{\sigma^2} \qquad \text{and} \qquad \rho = \sigma^{-2}, \qquad \text{or} \tag{5}$$

$$\mu = \frac{\tau}{\rho} \qquad \text{and} \qquad \sigma^2 = \rho^{-1}. \tag{6}$$

Also note that if $X \sim \mathcal{N}\left(\cdot;\mu,\sigma^2\right)$ we know that $\mathbb{E}\left[X\right] = \mu$, $\mathbb{E}\left[X^2\right] = \mu^2+\sigma^2$, and $\mathbb{V}\left[X\right] = \sigma^2$. An important quantity that we often use in relation to Gaussian densities is the cumulative distribution function (cdf) of the Gaussian density.

**Definition 2** (Gaussian Cumulative Distribution Function). Given parameters $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$, the cumulative distribution function $\Phi\left(\cdot;\mu,\sigma^2\right)$ of a Gaussian $\mathcal{N}\left(\cdot;\mu,\sigma^2\right)$ for any $t \in \mathbb{R}$ is defined by

$$\Phi\left(t;\mu,\sigma^2\right) := \int_{-\infty}^{t}\mathcal{N}\left(x;\mu,\sigma^2\right)\,\mathrm{d}x, \tag{7}$$

$$\Phi\left(t\right) := \Phi\left(t;0,1\right). \tag{8}$$

Due to the symmetry of the Gaussian density around $\mu$, the cumulative distribution function has some nice properties.

**Theorem 1** (Gaussian Cumulative Distribution Function). *Given parameters $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$, we have for any $t \in \mathbb{R}$*

$$\Phi\left(t;\mu,\sigma^2\right) = \Phi\left(\frac{t-\mu}{\sigma}\right), \tag{9}$$

$$\Phi\left(t\right) = 1-\Phi\left(-t\right). \tag{10}$$

# Truncated Gaussians

Before we prove the main result, we derive the following lemma using standard analysis results.

**Lemma 1** (Moments of a doubly-truncated Gaussian). *Given $l < u \in \mathbb{R}$, $\mu \in \mathbb{R}$ and $\sigma^2 \in \mathbb{R}^+$, we say that $X$ is distributed according to a double-truncated Gaussian if the density $p_X(\cdot)$ is proportional to*

$$p_X(x) \quad\propto\quad \mathbb{I}\left(l \le x < u\right)\cdot\mathcal{N}\left(x;\mu,\sigma^2\right).$$

*Then, we know that*

$$\mathbb{E}\left[X^0\right] = \Phi\left(u;\mu,\sigma^2\right) - \Phi\left(l;\mu,\sigma^2\right) , \tag{11}$$

$$\mathbb{E}\left[X^1\right] = \mu + \sigma^2 \cdot \frac{\mathcal{N}\left(l;\mu,\sigma^2\right) - \mathcal{N}\left(u;\mu,\sigma^2\right)}{\Phi\left(u;\mu,\sigma^2\right) - \Phi\left(l;\mu,\sigma^2\right)} , \tag{12}$$

$$\mathbb{E}\left[X^2\right] = \mu^2 + \sigma^2 \cdot \left(1 - \frac{(\mu+u)\cdot\mathcal{N}\left(u;\mu,\sigma^2\right) - (\mu+l)\cdot\mathcal{N}\left(l;\mu,\sigma^2\right)}{\Phi\left(u;\mu,\sigma^2\right) - \Phi\left(l;\mu,\sigma^2\right)}\right) . \tag{13}$$

*Proof.* To prove (11), we use Definition 2 to see that

$$\mathbb{E}\left[X^0\right] = \int_l^u \mathcal{N}\left(x;\mu,\sigma^2\right)\,\mathrm{d}x = \Phi\left(u;\mu,\sigma^2\right) - \Phi\left(l;\mu,\sigma^2\right) .$$

First note that using the chain rule of differentiation and (1)

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathcal{N}\left(x;\mu,\sigma^2\right) = \mathcal{N}\left(x;\mu,\sigma^2\right)\cdot\frac{\mathrm{d}}{\mathrm{d}x}\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) = -\left(\frac{x-\mu}{\sigma^2}\right)\cdot\mathcal{N}\left(x;\mu,\sigma^2\right) \tag{14}$$

To prove (12), we now use (14) and notice that

$$\frac{\mathrm{d}}{\mathrm{d}x} - \sigma^2\cdot\mathcal{N}\left(x;\mu,\sigma^2\right) = (x-\mu)\cdot\mathcal{N}\left(x;\mu,\sigma^2\right) .$$

Thus, we see that

$$\begin{aligned}
\mathbb{E}\left[X^1 - \mu\right] &= \frac{1}{\mathbb{E}\left[X^0\right]}\cdot\int_l^u (x-\mu)\cdot\mathcal{N}\left(x;\mu,\sigma^2\right)\,\mathrm{d}x \\
&= \frac{1}{\mathbb{E}\left[X^0\right]}\cdot\left[-\sigma^2\cdot\mathcal{N}\left(x;\mu,\sigma^2\right)\right]_l^u \\
&= \sigma^2\cdot\frac{\mathcal{N}\left(l;\mu,\sigma^2\right) - \mathcal{N}\left(u;\mu,\sigma^2\right)}{\Phi\left(u;\mu,\sigma^2\right) - \Phi\left(l;\mu,\sigma^2\right)} .
\end{aligned}$$

To prove (13), we use Definition 2 and (14) and notice that

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}x}\sigma^2\cdot\left[\Phi\left(x;\mu,\sigma^2\right) - (\mu+x)\cdot\mathcal{N}\left(x;\mu,\sigma^2\right)\right] \\
&= \sigma^2\cdot\left[\mathcal{N}\left(x;\mu,\sigma^2\right) - (\mu+x)\cdot\frac{\mathrm{d}\,\mathcal{N}\left(x;\mu,\sigma^2\right)}{\mathrm{d}x} - \frac{\mathrm{d}\ (\mu+x)}{\mathrm{d}x}\cdot\mathcal{N}\left(x;\mu,\sigma^2\right)\right] \\
&= \sigma^2\cdot\left[\mathcal{N}\left(x;\mu,\sigma^2\right) + \frac{(\mu+x)\cdot(x-\mu)}{\sigma^2}\cdot\mathcal{N}\left(x;\mu,\sigma^2\right) - \mathcal{N}\left(x;\mu,\sigma^2\right)\right] \\
&= (x+\mu)\cdot(x-\mu)\cdot\mathcal{N}\left(x;\mu,\sigma^2\right) \\
&= \left(x^2-\mu^2\right)\cdot\mathcal{N}\left(x;\mu,\sigma^2\right) .
\end{aligned}$$

Thus, we see that

$$\begin{aligned}
\mathbb{E}\left[X^2 - \mu^2\right] &= \frac{1}{\mathbb{E}\left[X^0\right]}\cdot\int_l^u \left(x^2-\mu^2\right)\cdot\mathcal{N}\left(x;\mu,\sigma^2\right)\,\mathrm{d}x \\
&= \frac{1}{\mathbb{E}\left[X^0\right]}\cdot\sigma^2\cdot\left[\Phi\left(x;\mu,\sigma^2\right) - (\mu+x)\cdot\mathcal{N}\left(x;\mu,\sigma^2\right)\right]_l^u \\
&= \sigma^2\cdot\frac{\Phi\left(u;\mu,\sigma^2\right) - (\mu+u)\cdot\mathcal{N}\left(u;\mu,\sigma^2\right) - \Phi\left(l;\mu,\sigma^2\right) + (\mu+l)\cdot\mathcal{N}\left(l;\mu,\sigma^2\right)}{\Phi\left(u;\mu,\sigma^2\right) - \Phi\left(l;\mu,\sigma^2\right)} \\
&= \sigma^2\cdot\left(1 - \frac{(\mu+u)\cdot\mathcal{N}\left(u;\mu,\sigma^2\right) - (\mu+l)\cdot\mathcal{N}\left(l;\mu,\sigma^2\right)}{\Phi\left(u;\mu,\sigma^2\right) - \Phi\left(l;\mu,\sigma^2\right)}\right) .
\end{aligned}$$

□

Equipped with this lemma, we can now prove the main result.

**Theorem 2** (Mean and Variance of a doubly-truncated Gaussian). *Given $l < u \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\sigma^2 \in \mathbb{R}^+$ and a doubly-truncated Gaussian with density $p_X(\cdot)$*

$$p_X(x) \quad \propto \quad \mathbb{I}\left(l \le x < u\right) \cdot \mathcal{N}\left(x; \mu, \sigma^2\right) ,$$

*we know that*

$$\mathbb{E}\left[X\right] = \mu + \sigma \cdot v_{\frac{l}{\sigma},\frac{u}{\sigma}}\left(\frac{\mu}{\sigma}\right) , \tag{15}$$

$$\mathbb{V}\left[X\right] = \sigma^2 \cdot \left[1 - w_{\frac{l}{\sigma},\frac{u}{\sigma}}\left(\frac{\mu}{\sigma}\right)\right] , \tag{16}$$

*where*

$$v_{l,u}\left(t\right) = \frac{\mathcal{N}\left(l-t\right) - \mathcal{N}\left(u-t\right)}{\Phi\left(u-t\right) - \Phi\left(l-t\right)} \tag{17}$$

$$w_{l,u}\left(t\right) = \frac{\left(u+t\right) \cdot \mathcal{N}\left(u-t\right) - \left(l+t\right) \cdot \mathcal{N}\left(l-t\right)}{\Phi\left(u-t\right) - \Phi\left(l-t\right)} + v_{l,u}\left(t\right) \cdot \left[2t + v_{l,u}\left(t\right)\right] . \tag{18}$$

*Proof.* The first result (15) follows directly from (12) in Lemma 1 using

$$\mathcal{N}\left(x; \mu, \sigma^2\right) = \frac{1}{\sigma} \cdot \mathcal{N}\left(\frac{x-\mu}{\sigma}\right) .$$

The second result (16) follows from (13) in Lemma 1 noting the variance decomposition theorem:

$$\begin{aligned}
\mathbb{V}\left[X\right] &= \mathbb{E}\left[X^2\right] - \left(\mathbb{E}\left[X\right]\right)^2 \\
&= \mu^2 + \sigma^2 \cdot \left(1 - \underbrace{\frac{\left(\frac{\mu}{\sigma} + \frac{u}{\sigma}\right) \cdot \mathcal{N}\left(\frac{u}{\sigma} - \frac{\mu}{\sigma}\right) - \left(\frac{\mu}{\sigma} + \frac{l}{\sigma}\right) \cdot \mathcal{N}\left(\frac{l}{\sigma} - \frac{\mu}{\sigma}\right)}{\Phi\left(\frac{u}{\sigma} - \frac{\mu}{\sigma}\right) - \Phi\left(\frac{l}{\sigma} - \frac{\mu}{\sigma}\right)}}_{A_{\frac{l}{\sigma},\frac{u}{\sigma},\frac{\mu}{\sigma}}}\right) - \left[\mu + \sigma \cdot v_{\frac{l}{\sigma},\frac{u}{\sigma}}\left(\frac{\mu}{\sigma}\right)\right]^2 \\
&= \sigma^2 \cdot \left(1 - A_{\frac{l}{\sigma},\frac{u}{\sigma},\frac{\mu}{\sigma}}\right) - 2 \cdot \mu\sigma \cdot v_{\frac{l}{\sigma},\frac{u}{\sigma}}\left(\frac{\mu}{\sigma}\right) - \sigma^2 \cdot v^2_{\frac{l}{\sigma},\frac{u}{\sigma}}\left(\frac{\mu}{\sigma}\right) \\
&= \sigma^2 \cdot \left(1 - A_{\frac{l}{\sigma},\frac{u}{\sigma},\frac{\mu}{\sigma}} - 2 \cdot \frac{\mu}{\sigma} \cdot v_{\frac{l}{\sigma},\frac{u}{\sigma}}\left(\frac{\mu}{\sigma}\right) - v^2_{\frac{l}{\sigma},\frac{u}{\sigma}}\left(\frac{\mu}{\sigma}\right)\right) \\
&= \sigma^2 \cdot \left(1 - \left(A_{\frac{l}{\sigma},\frac{u}{\sigma},\frac{\mu}{\sigma}} + v_{\frac{l}{\sigma},\frac{u}{\sigma}}\left(\frac{\mu}{\sigma}\right) \cdot \left[2 \cdot \frac{\mu}{\sigma} + v_{\frac{l}{\sigma},\frac{u}{\sigma}}\left(\frac{\mu}{\sigma}\right)\right]\right)\right) \\
&= \sigma^2 \cdot \left(1 - w_{\frac{l}{\sigma},\frac{u}{\sigma}}\left(\frac{\mu}{\sigma}\right)\right)
\end{aligned}$$

□