1 The Skew-Normal

The skew-normal distribution is similar to the normal but with an added parameter for skew. In this section, we'll introduce

1.1 Basics

Let Y be a skew-normal distribution, with location parameter $\mu \in \mathbb{R}$, scale parameter $\sigma > 0$, and shape parameter $\lambda \in \mathbb{R}$, denoted $SN(\mu, \sigma, \lambda)$. Then Y has pdf

$$f_Y(x) = \frac{2}{\sigma} \cdot \phi\left(\frac{x-\mu}{\sigma}\right) \cdot \Phi\left(\frac{\lambda(x-\mu)}{\sigma}\right) \tag{1}$$

where ϕ is the standard normal pdf and Φ is the standard normal cdf. Some other basic properties of Y, given by Pewsey (2000), are

$$\begin{aligned}
E(Y) &= \mu + b\delta\sigma \\
E(Y^2) &= \mu^2 + 2b\delta\mu\sigma + \sigma^2 \\
Var(Y) &= \sigma^2(1 - b^2\delta^2) \\
E(Y^3) &= \mu^3 + 3b\delta\mu^2\sigma + 3\mu\sigma^2 + 3b\delta\sigma^3 - b\delta^3\sigma^3
\end{aligned} \tag{2}$$

where $b = \sqrt{\frac{2}{\pi}}$ and $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$.

The $SN(0, 1, \lambda)$ distribution is called the standard skew-normal, and has pdf

$$f_z(x) = 2 \cdot \phi(x) \cdot \Phi(-\lambda x) \tag{3}$$

Like the normal and standard normal, you can arrive at the standard skew normal by applying the transformation $\frac{Y-\mu}{\sigma}$ to the skew-normal distribution.

A natural question to ask is how the skew-normal relates to the normal. Fortunately, the connection is very intuitive: Letting $\lambda = 0$ in (1) returns us to the normal pdf:

$$
\begin{aligned}
f_Y(x|\lambda=0) &= \frac{2}{\sigma}\cdot\phi\left(\frac{x-\mu}{\sigma}\right)\Phi(0)\\
&= \frac{2}{\sigma}\cdot\phi\left(\frac{x-\mu}{\sigma}\right)\cdot 0.5\\
&= \frac{1}{\sigma}\cdot\phi\left(\frac{x-\mu}{\sigma}\right)\\
&= \frac{1}{\sigma}\cdot\frac{1}{\sqrt{2\pi}}\ \cdot\ \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\\
&= \frac{1}{\sqrt{2\pi\sigma^2}}\ \cdot\ \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)
\end{aligned}
$$

1.2 Four Properties

Chang et al. (2008) gives four handy properties of the skew-normal distribution:

Property 1. *If $Z \sim SN(0,1,\lambda)$, then $(-Z) \sim SN(0,1,-\lambda)$.*

Proof. The standard normal pdf is an even function: $\phi(-x) = \frac{1}{\sqrt{2\pi}}\ e^{-(-x)^2/2} = \frac{1}{\sqrt{2\pi}}\ e^{-x^2/2} = \phi(x)$. But the standard normal cdf, $\Phi(x) = \int_{-\infty}^{\infty}\phi(x),$ is not, being 0 near $-\infty$ and 1 near ∞. Thus,

$$
\begin{aligned}
f_{(-Z)}(x) &= f_Z(-x)\\
&= 2\cdot\phi(-x)\cdot\Phi(-\lambda x)\\
&= 2\cdot\phi(x)\cdot\Phi(-\lambda x)
\end{aligned}
$$

which is the pdf of $SN(0,1,-\lambda)$. $\mathcal{Q.E.D.}$

Property 2. *As $\lambda \to \pm\infty$, $SN(0,1,\lambda)$ tends to the half normal distribution, $\pm|N(0,1)|$.*

To prove our theorem, it is helpful to formally define the half normal distribution:

Lemma 2.1. *Let $X \sim N(0,\sigma^2)$. Then the distribution of $|X|$ is a half-normal random variable with parameter σ and*

$$
f_{|X|}(x) = \begin{cases} 2\cdot f_X(x) & \textit{when } 0 < x < \infty\\ 0 & \textit{everywhere else}\end{cases}
$$

Proof. Let $X \sim N(0,\sigma^2)$, defined over $A = (-\infty,\infty)$. Define

$$Y = |X| = \begin{cases} -x & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ x & \text{if } x > 0 \end{cases}$$

Y is not one-to-one over A. However, we can partition A into disjoint subsets $A_1 = (-\infty, 0)$, $A_2 = (0, \infty)$, and $A_3 = \{0\}$ such that $A = A_1 \cup A_2 \cup A_3$ and Y is one-to-one over each A_i. We can then transform each piece separately using Theorem 6.3.2 from Bain and Engelhardt (1992):

On A_1: $y = -x \longrightarrow x = -y$ and $\mathbb{J} = \left|\frac{dx}{dy}\right| = |-1| = 1$, yielding

$$\begin{aligned} f_Y(y) &= f_X(x) \cdot \mathbb{J} \\ &= f_X(-y) \cdot 1 \\ &= \frac{1}{\sqrt{2\pi\sigma}}\, e^{-\frac{(-y)^2}{2\sigma}} \\ &= \frac{1}{\sqrt{2\pi\sigma}}\, e^{-\frac{y^2}{2\sigma}} \\ &= f_X(y) \end{aligned} \tag{4}$$

over the domain $A_1 : -\infty < x < 0 \longrightarrow -\infty < -y < 0 \longrightarrow 0 < y < \infty : B_1$.

Similarly, on A_2: $y = x \longrightarrow x = y$ and $\mathbb{J} = \left|\frac{dx}{dy}\right| = |1| = 1$, yielding

$$\begin{aligned} f_Y(y) &= f_X(x) \cdot \mathbb{J} \\ &= f_X(y) \cdot 1 \\ &= f_X(y) \end{aligned} \tag{5}$$

over the domain $A_2 : 0 < x < \infty \longrightarrow 0 < y < \infty : B_2$.

On A_3, we have $x = 0, y = 0$ and $\mathbb{J} = \left|\frac{dx}{dy}\right| = |0| = 0$, yielding $f_Y(y) = f_X(x) \cdot \mathbb{J} = f_X(x) \cdot 0 = 0$.

Now, by Theorem 6.3.10 from Bain and Engelhardt (1992), we achieve our result by simply summing (4) and (5).

$$\begin{aligned} f_Y(y) &= \{f_Y(y) \text{ over } A_1\} + \{f_Y(y) \text{ over } A_2\} \\ &= f_X(y) + f_X(y) \\ &= 2 \cdot f_X(y) \end{aligned} \tag{6}$$

over $B = B_1 \cup B_2 = (0, \infty)$, and 0 otherwise. $\mathcal{Q.E.D.}$

With Lemma 2.1, we can easily show our property:

Proof of Property 2. Let $Z \sim SN(0,1,\lambda)$. Recall that $f_z(x) = 2 \cdot \phi(x) \cdot \Phi(\lambda x)$.

Consider $\lim_{\lambda \to \infty} f_X(x)$. When x is negative, $\lambda x \to -\infty$ and thus $\Phi(\lambda x) \to 0$. When x is positive, however, $\lambda x \to \infty$ and $\Phi(\lambda x) \to 1$. Thus

$$\lim_{\lambda \to \infty} 2 \cdot \phi(x) \cdot \Phi(\lambda x) \;=\; \begin{cases} 0 & \text{when } x \le 0 \\ 2 \cdot \phi(x) & \text{when } x > 0 \end{cases} \;=\; |N(0,1)| \tag{7}$$

In $\lim_{\lambda \to -\infty} f_X(x)$, the signs are reversed. When x is negative, $\lambda x \to \infty$ and $\Phi(\lambda x) \to 1$. When x is positive, $\lambda x \to -\infty$ and $\Phi(\lambda x) \to 0$. Thus,

$$\lim_{\lambda \to -\infty} 2 \cdot \phi(x) \cdot \Phi(\lambda x) \;=\; \begin{cases} 2 \cdot \phi(x) & \text{when } x < 0 \\ 0 & \text{when } x \ge 0 \end{cases} \;=\; -|N(0,1)| \tag{8}$$

$\mathcal{Q.E.D.}$

Property 3. *If $Z \sim SN(0,1,\lambda)$, then $Z^2 \sim \chi^2_1$ (chi-square with 1 degree of freedom).*

Proof. To prove our result, we make use of Lemma 1 in Azzalini (2005):

Lemma 3.1. *If f_0 is a one-dimensional probability density function symmetric about 0, and G is a one-dimensional distribution function such that G' exists and is a density symmetric about 0, then*

$$f(z) = 2 \cdot f_0(z) \cdot G\{w(z)\} \quad (-\infty < z < \infty) \tag{9}$$

is a density function for any odd function $w(\cdot)$.

Notice that $\phi(x)$ is a one-dimensional probability density function symmetric about 0, and $\Phi(x)$ is a one-dimensional distribution function such that Φ' exists and is a density symmetric about 0. Furthermore, λx is an odd function. Thus, $f_z(z) = 2 \cdot \phi(z) \cdot \Phi(\lambda z)$ conforms to equation (9). With that in mind, the corollary to this lemma provides a very useful result:

Corollary 3.1 (Perturbation Invariance)**.** *If $Y \sim f_0$ and $Z \sim f$, then $|Y| \overset{d}{=} |Z|$, where the notation $\overset{d}{=}$ denotes equality in distribution.*

Thus, we can treat ϕ and Z as being equal in distribution. We will now show that $\phi^2 \sim \chi^2_1$:

$$
\begin{aligned}
M_{\phi^2}(t) &= E[e^{tx^2}] \\
&= \int_{-\infty}^{\infty} e^{tx^2} \left[\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \right] dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{tx^2 - x^2/2}\, dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}(1-2t)}\, dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\sqrt{1-2t}\, x)^2}\, dx
\end{aligned}
$$

Let $u = (\sqrt{1-2t})\, x$; then we have $du = \sqrt{1-2t}$, $dx = \frac{du}{\sqrt{1-2t}}$, and our limits become $x \to -\infty \Rightarrow u \to -\infty$ and $x \to \infty \Rightarrow u \to \infty$.

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \left(\frac{1}{\sqrt{1-2t}} \right) du \\
&= \frac{1}{\sqrt{1-2t}} \underbrace{\left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du \right)}_{\phi(u) \text{ integrated over } (-\infty,\infty) = 1} \\
&= \frac{1}{\sqrt{1-2t}}
\end{aligned}
$$

which is the MGF of the χ^2_1. Since Z is equal in distribution to ϕ, we can also conclude that $Z^2 \sim \chi^2_1$. $\mathcal{Q.E.D.}$

Property 4. *The MGF of $SN(0, 1, \lambda)$ is*

$$M(t|\lambda) = 2 \cdot \Phi(\delta t) \cdot e^{t^2/2} \tag{10}$$

where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$ and $t \in (-\infty, \infty)$.

Proof. According to Equation 5 in Azzalini (2005), the MGF of $SN(\mu, \sigma^2, \lambda)$ is

$$M(t) \;=\; E\{e^{tY}\} \;=\; 2 \cdot \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \cdot \Phi(\delta \sigma t)$$

where $\delta = \frac{\lambda}{1+\lambda^2} \in (-1, 1)$. It follows that the MGF of the $SN(0, 1, \lambda)$ is

$$2 \cdot \exp\left(0 \cdot t + \frac{1 \cdot t^2}{2}\right) \cdot \Phi(\delta \cdot 1 \cdot t) \;=\; 2 \cdot e^{t^2/2} \cdot \Phi(\delta t)$$

$\mathcal{Q.E.D.}$

References

Adelchi Azzalini. The skew-normal distribution and related multivariate families. *Scandinavian Journal of Statistics*, 32:159–188, 2005. Advanced properties of the skew-normal.

Lee J. Bain and Max Engelhardt. *Introduction to Probability and Mathematical Statistics*. Duxbury, 2nd edition, 1992. Stat 570/571 textbook.

Ching-Hui Chang, Jyh-Jiuan Lin, Nabendu Pal, and Miao-Chen Chiang. A note on improved approximation of the binomial distribution by the skew-normal distribution. *American Statistical Association*, 62(2):167–170, May 2008. Main article.

Arthur Pewsey. Problems of inference for azzalini's skew-normal distribution. *Journal of Applied Statistics*, 27(7):859–870, 2000. Mean, variance, expected value of the skew normal.