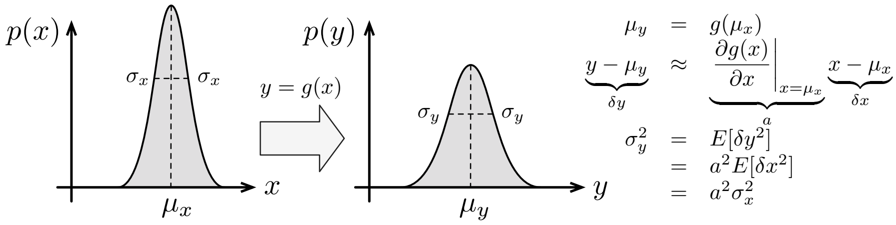


**Figure 2.5** Passing a one-dimensional Gaussian through a deterministic nonlinear function, $g(\cdot)$. Here we linearize the nonlinearity in order to propagate the variance approximately.

generated by sampling $x$ a large number of times and passing these through the nonlinearity individually, then binning. These approaches agree very well, validating our method of changing variables.

Note that $p(y)$ is no longer Gaussian owing to the nonlinear change of variables. We can verify numerically that the area under this function is indeed 1 (i.e., it is a valid PDF). It is worth noting that had we not been careful about handling the change of variables and including the $\frac{1}{y}$ factor, we would not have a valid PDF.

### *General Case via Linearization*

Unfortunately, (2.76) cannot be computed in closed form for every $\mathbf{g}(\cdot)$ and becomes more difficult in the multivariate case than the scalar one. Moreover, when the nonlinearity is stochastic (i.e., $\mathbf{R} > 0$), our mapping will never be invertible due to the extra input coming from the noise, so we need a different way to transform our Gaussian. There are several different ways to do this, and in this section, we look at the most common one, *linearization*.

We linearize the nonlinear map such that

$$\begin{aligned} \mathbf{g}(\mathbf{x}) &\approx \boldsymbol{\mu}_y + \mathbf{G}(\mathbf{x} - \boldsymbol{\mu}_x), \\ \mathbf{G} &= \left.\frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\boldsymbol{\mu}_x}, \\ \boldsymbol{\mu}_y &= \mathbf{g}(\boldsymbol{\mu}_x), \end{aligned} \tag{2.85}$$

where $\mathbf{G}$ is the Jacobian of $\mathbf{g}(\cdot)$, with respect to $\mathbf{x}$. This allows us to then pass the Gaussian through the linearized function in closed form; it is an approximation that works well for mildly nonlinear maps.

Figure 2.5 depicts the process of passing a one-dimensional Gaussian PDF through a deterministic nonlinear function, $g(\cdot)$, that has been linearized. In general, we will be making an inference though a stochastic function, one that introduces additional noise.