Delta Method: The Importance of Var(X)

This simple example was discussed in the February 13, 2007 class: Consider a random variable X, with mean given by $E(X)=\mu_X$ and variance given by $Var(X)=\sigma_X^2$, and suppose we are interested in the moments of the new random variable $Y=-\log(X)$. Then the delta method yields:

a) $E(Y)\cong -\log(\mu_X)$ as a first order approximation,

b) $E(Y)\cong -\log(\mu_X)+\sigma_X^2/2\mu_X^2$ as a second order approximation,

c) $Var(Y)\cong \sigma_X^2/\mu_X^2$ as a first order approximation.

Now consider the specific example of X having density function $f_X(x)=(\theta+1)x^\theta$ for $0<x<1$ (with $\theta\ge 0$). It is easy to calculate $\mu_X=(\theta+1)/(\theta+2)$ and $\sigma_X^2=(\theta+1)/(\theta+2)^2(\theta+3)$. Note, in particular, that as θ gets larger, μ_X gets closer and closer to 1 and σ_X^2 gets smaller and smaller.

For this specific example, the delta method approximations become:

a) $E(Y)\cong -\log\big((\theta+1)/(\theta+2)\big)$,

b) $E(Y)\cong -\log\big((\theta+1)/(\theta+2)\big)+1/2(\theta+1)(\theta+3)$,

c) $Var(Y)\cong 1/(\theta+1)(\theta+3)$

But, for this specific example, it is also easy to calculate the exact values of $E(Y)$ and $Var(Y)$. To evaluate $E(Y^k)=\int_0^1(-\log x)^k f_X(x)dx$, transform to $y=-\log x$ so the integrand becomes a Gamma density. You find $E(Y^k)=\Gamma(k+1)/(\theta+1)^k$ ➔ $E(Y)=1/(\theta+1)$ and $Var(Y)=1/(\theta+1)^2$.

Comparing the exact values and delta method approximations (RE = relative error, expressed as %):

θ	σ_X^2	$E(Y)$	a) (RE)	b) (RE)	$Var(Y)$	c) (RE)	$SD(Y)$	c) (RE)
0	0.083	1.000	0.693 (-31)	0.859 (-14)	1.000	0.333 (-67)	1.000	0.577 (-42)
1	0.056	0.500	0.406 (-19)	0.468 (-6.4)	0.250	0.125 (-50)	0.500	0.354 (-29)
2	0.038	0.333	0.288 (-14)	0.321 (-3.7)	0.111	0.067 (-40)	0.333	0.258 (-23)
3	0.027	0.250	0.223 (-11)	0.244 (-2.4)	0.063	0.042 (-33)	0.250	0.204 (-18)
4	0.020	0.200	0.182 (-8.8)	0.197 (-1.7)	0.040	0.029 (-29)	0.200	0.169 (-16)
5	0.015	0.167	0.154 (-7.5)	0.165 (-1.3)	0.028	0.021 (-25)	0.167	0.144 (-13)
10	0.006	0.091	0.087 (-4.3)	0.091 (-0.4)	0.008	0.007 (-15)	0.091	0.084 (-8.0)
100	$9\text{x}10^{-5}$	0.00990	0.00985 (-0.5)	0.00990 (-0.006)	$9.8\text{x}10^{-5}$	$9.6\text{x}10^{-5}$ (-1.9)	$9.9\text{x}10^{-3}$	$9.8\text{x}10^{-3}$ (-1.0)

Note how the RE in each approximation decreases as θ increases (i.e., as σ_X^2 increases). Also note the substantial improvement provided by the 2nd order approximation to $E(Y)$ even for smaller values of θ. For this example, the value of θ has to get pretty large (the value of σ_X^2 has to get pretty small) before the approximations to $Var(Y)$ and $SD(Y)$ become reasonably accurate. You might want to plot the density functions $f_X(x)$ for different values of θ to see how much those distributions change as θ changes.