Question 2 Find the smallest double precision number point number $x > 16$ (i.e., the next double precision number after 16), and the largest double precision number $x < 16$ (i.e., the previous double precision number before 16). In both cases, find all bits of x and the difference from 16, and compare the difference to the machine epsilon eps.

Because $16 = 1.0000000000 \ldots\ldots 00000 \times 2^4$

$\boldsymbol{x > 16}$: $x = 1.0000000000 \ldots\ldots 00001 \times 2^4 = 16 + 2^{-52} \times 2^4 = 16 + 2^{-48}$

$1.0000000000 \ldots\ldots 00001$ ($52 bits\ after$ ".")

$x - 16 = 2^{-48} > \epsilon = 2^{-52}$

Because $15 = 1.1110000000 \ldots 000000 \times 2^3$

$1.1111111111111 \ldots 111111$ ($52 bits\ after$ ".")

$\boldsymbol{x < 16}$: $x = 1.1111111111111 \ldots 111111 \times 2^3$

$x - 16 = 1.0000000000 \ldots\ldots 00000 \times 2^4$

$- 1.1111111111111 \ldots 111111 \times 2^3 = 2^{-49} > \epsilon = 2^{-52}$

Question 3 Find the forward and backward error for the following functions, where the root is $\frac{1}{3}$, and the approximate root is $x_a = 0.3333$:

(a) $f(x) = 3x - 1$

$$FE = |r - x_a| = 3.3333 \times 10^{-5}$$

$$BE = |3 \times 0.3333 - 1| = 1 \times 10^{-4}$$

(b) $f(x) = (3x - 1)^2$

$$FE = |r - x_a| = 3.3333 \times 10^{-5}$$

$$BE = |(3 \times 0.3333 - 1)^2| = 1 \times 10^{-8}$$

(c) $f\,(x) = (3x\ - 1)^3$

$$FE = |r - x_a| = 3.3333 \times 10^{-5}$$

$$BE = f\,(0.3333) = |(3 \times 0.3333\ - 1)^3| = 1 \times 10^{-12}$$

(d) $f\,(x) = (3x\ - 1)^{\frac{1}{3}}$

$$FE = |r - x_a| = 3.3333 \times 10^{-5}$$

$$BE = f\,(0.3333) = |(3 \times 0.3333\ - 1)^{\frac{1}{3}}| = 0.0464$$