$(x-1)^2 e^x$ $\frac{d}{dx}(x-1)^2 = 2x-2$ $\frac{d}{dx} e^x = e^x$ $(2x-2)e^x + e^x(x-1)^2$ $e^x((2x-2)+(x^2-2x+1)) f'(x) = e^x(x^2-1)$

$f(x) = f(P_0) + f'(P_0)(x-P_0) + \frac{f''(P_0)}{2}(x-P_0)^2 + \dots$ $f(x) = f(P_0) + f'(P_0)(x-P_0) + \frac{f''(\eta(P_0))}{2}(x-P_0)^2$

$f(x)$ @ $x=p$ $f(x) = f(P_0) + f'(P_0)(P-P_0) + \frac{f''(\eta(P_0))}{2}(P-P_0)^2$ $f'(P_0)(P-P_0) \ll \frac{f''(\eta(P_0))}{2}(P-P_0)^2$

$0 = f(P_0) + f'(P_0)(P-P_0)$ $-f(P_0) = f'(P_0)(P-P_0)$ $P - P_0 = -\frac{f(P_0)}{f'(P_0)}$

$P = P_0 - \frac{f(P_0)}{f'(P_0)}$

$g(x) = x - \frac{\cancel{e^x}(x-1)^2}{\cancel{e^x}(x^2-1)} = g(x) = x - \frac{(x-1)^2}{(x^2-1)}$ $f'(1) = 0$ so $x^* = 1$

$g(x) = x - \frac{(x-1)\cancel{(x-1)}}{(x+1)\cancel{(x-1)}} = x - \frac{x-1}{x+1} = \frac{x^2+x}{x+1} - \frac{x-1}{x+1} = \frac{x^2+1}{x+1}$

$\lim_{x_n \to x^*} \frac{\left|\frac{x_n^2+1}{x_n+1} - x^*\right|}{|x_n - x^*|}$ $\frac{\left|\frac{x_n^2+1}{x_n+1} - 1\right|}{|x_n - 1|} = \frac{\frac{x_n^2+1}{x_n+1} - \frac{x_n+1}{x_n+1}}{|x_n - 1|}$ $\frac{\frac{x_n^2 - x_n}{x_n+1}}{x_n - 1}$

$x^* = 1$

$\frac{x_n(x_n-1)}{(x_n+1)(x_n-1)} = \frac{x_n}{x_n+1}$ $\lim_{x_n \to 1} \frac{x_n}{x_n+1} = \boxed{\frac{1}{2}}$

b.) implementing this in C++ at $x_0 = 2$ it took 28 iterations to reach a tolerance of 10^{-8}

C.) it would be easy to implement bisection, but not easier than newton's method after simplication of the function

1.) bisection is not efficient, however it is fairly robust given the function doesn't bounce of the axis. It only requires a positive point, and a negative point, but other knowledge isn't strictly necessary. It does not require any addition knowledge over a functions derivatives. This function does generalize easily.
Newton's method is efficient, it is mostly robust with few situations where it doesn't converge. It require knowledge of whether it is $C^2[a,b]$ and the derivative at the intial guess. It does require f to statisfy minimum smoothness. It can generalize to multi variables
Secant method is mostly efficient, and is not as robust as bisection in that it doesn't always converge. It requires 2 intial points for it to function. It doesn't required f to statisfy minimum smoothness. It seems like the method can be generalized to multi variable functions.

2.) One advantage is that it converges on the root fast, and in addition it is not bad when it comes to computational efficiency. One big disadvangtage is that it requries $C^2[a,b]$ and for $f'(x)$ to be non zero. It Also only converges if $0 \leq g'(x) < 1$

3.) The biggest disadvantage of secant of newton's method is the rate of convergence is worse for secant. The biggest advantage is the lact of derivative knowledge

4.) $$f(x) = f(p^*) + f'(p^*)(x-p^*) + \frac{f''(p^*)}{2!}(x-p^*)^2 + \frac{f'''(p^*)}{3!}(x-p^*)^3 + \dots$$

$\eta(p^*) \in (x, p^*)$ $$f(x) = f(p^*) + f'(p^*)(x-p^*) + \frac{f''(p^*)}{2!}(x-p^*)^2 + \frac{f'''(\eta(p^*))}{3!}(x-p^*)^3$$ $f(x)$ $x = p$

$$f(p) = 0 = f(p^*) + f'(p^*)(p-p^*) + \frac{f''(p^*)}{2!}(p-p^*)^2 + \frac{f'''(\eta(p^*))}{3!}(p-p^*)^3$$ $|p-p^*|^3 \ll |p-p^*|^2$

$$0 = f(p^*) + f'(p^*)(p-p^*) + \frac{f''(p^*)}{2!}(p-p^*)^2$$

$a = \frac{f''(p^*)}{2!}$ $z = (p-p^*)$
$b = f'(p^*)$
$c = f(p^*)$

$$0 = c + bz + az$$

$$z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \qquad (p-p^*) = \frac{-f'(p^*) \pm \sqrt{f'(p^*)^2 - 4\left(\frac{f''(p^*)}{2}\right)(f(p^*))}}{-2\left(\frac{f''(p^*)}{2}\right)}$$

$$p = p^* + \frac{-f'(p^*) \pm \sqrt{f'(p^*)^2 - 2f''(p^*)f(p^*)}}{f''(p^*)}$$

$$p = p^* - \frac{f'(p^*) + \sqrt{f'(p^*)^2 - 2f''(p^*)f(p^*)}}{f''(p^*)} \quad \text{or} \quad p = p^* - \frac{f'(p^*) - \sqrt{f'(p^*)^2 - 2f''(p^*)f(p^*)}}{f''(p^*)}$$

given $f''(p^*) > 0$ and $f'(p^*)^2 - 2f''(p^*)f(p^*) \geq 0$

$$g(p_n) \leq g(p) + g'(p)(p_n - p) + \frac{g''(p)}{2!}(p_n - p)^2 + \frac{m}{6}(p_n - p)^3$$

$m \geq |g'''(x)|$ on (a, b)

$$g(p_n) - g(p) \approx |p_{n+1} - p| \leq \frac{m}{6}|p_n - p|^3$$

The biggest reason why it may not be used is the limited functions which are suitable for the method. The other is it requires much more knowledge than newton's which may be able to achieve the same hveristic time due to fewer operations of the functio