Problem: UCB and KL-UCB for Bernoulli Bandits,
prove that if $0 < \delta(t) \leq \delta'(t)$, then $ucb_a^t \leq ucb_a^{kl,t}$ for any number of pulls u_a^t and empirical mean $\hat{p}_a^t$

Solution:
Given,
UCB:

$$ucb_a^t = \hat{p}_a^t + \sqrt{\frac{1}{2u_a^t}\ln(\frac{1}{\delta(t)})}$$

KL-UCB:

$$ucb_a^{kl,t} = argmax_{q\in[0,1]} u_a^t KL(\hat{p}_a^t, q) = \ln(\frac{1}{\delta'(t)})$$

where, $\delta'(t) = 1/(t \ln t^c)$ for $c \geq 3$

To Prove:
if $0 \leq \delta'(t)' \leq \delta(t)$ then $ucb_a^t \leq ucb_a^{kl,t}$
for any number of pulls ucb_a^t and empirical mean $\hat{p}_a^t$

Proof:

$$u_a^t KL(\hat{p}_a^t, ucb_a^{kl,t}) = \ln(\frac{1}{\delta'(t)})$$

Using Pinsker's Inequality $[KL(x,y) \geq 2(x-y)^2]$,

$$u_a^t * 2(\hat{p}_a^t - ucb_a^{kl,t})^2 \leq u_a^t KL(\hat{p}_a^t, ucb_a^{kl,t}) = \ln\frac{1}{\delta'(t)}$$

$$(\hat{p}_a^t - ucb_a^{kl,t})^2 \leq \frac{1}{2u_a^t}\ln(\frac{1}{\delta'(t)})$$

$$|\hat{p}_a^t - ucb_a^{kl,t}| \leq \sqrt{\frac{1}{2u_a^t}\ln(\frac{1}{\delta'(t)})}$$

- Consider case 1, $\hat{p}_a^t \geq ucb_a^{kl,t}$:

$$\hat{p}_a^t - ucb_a^{kl,t} \leq \sqrt{\frac{1}{2u_a^t}ln(\frac{1}{\delta'(t)})}$$

$$ucb_a^{kl,t} \geq \hat{p}_a^t - \sqrt{\frac{1}{2u_a^t}ln(\frac{1}{\delta'(t)})}$$

Since $0 < \delta(t) \leq \delta'(t)$,
$\ln(\frac{1}{\delta(t)}) \leq \ln(\frac{1}{\delta'(t)})$

Therefore,
$ucb_a^{kl,t} \geq \hat{p}_a^t - \sqrt{\frac{1}{2u_a^t} ln(\frac{1}{\delta(t)})} = ucb_a^t$

- Consider case 2, $\hat{p}_a^t < ucb_a^{kl,t}$:

$$ucb_a^{kl,t} - \hat{p}_a^t \leq \sqrt{\frac{1}{2u_a^t} ln(\frac{1}{\delta'(t)})}$$

$$ucb_a^{kl,t} \leq \hat{p}_a^t + \sqrt{\frac{1}{2u_a^t} ln(\frac{1}{\delta'(t)})} \leq \hat{p}_a^t + \sqrt{\frac{1}{2u_a^t} ln(\frac{1}{\delta(t)})} = ucb_a^t$$

Since $0 < \delta(t) \leq \delta'(t)$,
$\ln(\frac{1}{\delta(t)}) \leq \ln(\frac{1}{\delta'(t)})$
Therefore,
$ucb_a^{kl,t} \geq \hat{p}_a^t - \sqrt{\frac{1}{2u_a^t} ln(\frac{1}{\delta(t)})} = ucb_a^t$

In both cases, we have shown that $ucb_a^t \leq ucb_a^{kl,t}$