Midterm questions

1. CIA doesn't hold in Bayesian learning, because unobservable state (f.e. number of succeses in the problem from problem set) depends on the previous unobservable state. However, even though theoretically CCP are no defined for such case, hazard rates calculated from data are estimates of CCP (MLE then chooses parameters so that model predicts hazard rates in data). So then the problem with Hotz-Miller approach is that inversion theorem doesn't work (due to cdf of errors $G(\varepsilon|x)$ not existing?

2. In the end of lecture 3 we claim that without loss of generality $u(x,\varepsilon) = u^*(x) + \varepsilon^*$. However, it implies that $\mathbb{E}(\varepsilon^*|x) = 0$. Then it is not the same as assuming $u(x,\varepsilon) = u(x) + \varepsilon$, (what we did in the beginning in lecture 6).

3. In lecture 6, slides 10, 12. What does it mean that "(x_t, d_t) is the DGP"? Same as model correctly specified? (since model generated data)

4. Lecture 7:

   - Example (slide 9). What does this last sentence mean? What are those differences: $u^*_{21}(1) - u_{21}(1)$ or $u^*_{21}(1) - u^*_{11}(1)$ or $u_{21}(1) - u_{11}(1)$? Why is not knowing c is important to identification? What if we knew it, are we able to identify anything? We still don't know $u^*_{21}(1)$, $u_{21}(1)$ or $u_{11}(1)$. Could you go over the example of smoking again?

   - Let's look at multinomial logit model. Parameters are not identified, unless we have a normalization. It seems we never view it as a problem, since differences between all options and a normalized option are identified. However, in long panels with no known payoff, the problem is similar: for given normalization we have a unique solution of all other payoffs, hence differences are identified up to normalization. Isn't it semantics that we call this model unidentified in primitives? Or do we say that multinomial logit also doesn't have identification in primitives, because of existence of observationally equivalent utilites (but then, is identification of primitives a big deal)? Or is the problem that in multinomial logit normalization is innocuous, since it doesn't affect estimates, while for long panels we need to normalize an action at eact period/state, so we impose structure across these normalized payoffs, which is additional assumption of the model?

   - Is first theorem good result or bad result for identification? Do we use it anywhere next? (Because we can derive Theorem on slide 21 and Corollary from slide 24 from Representation Theorem)

5. Lecture 9 – is ρ period dependence the same thing as a finite dependence? Is the use of finite dependence in CCP estimation just that we can sum things up to ρ instead of T? (It seems that we need to calculate ω weights, and need to assume finite dependence exists for any pair of choices).

see more questions below typos

Typos

1. Lecture 1, slide 8 – sum over z_{t+1} instead of z
2. Lecture 2, slide 10 – distribution of $(x_{m,t+1} - \gamma)/\sigma \sim N(0, (\alpha + t)^{-1} + 1)$
3. Lecture 2, slide 15 – previous distribution shows up in the first line, so if there was a typo, this also needs to be corrected.
4. Lecture 3, slide 10 – product of first term should be $t = 1$ up to $T - 1$
5. Lecture 3, slide 11 – same with first part (product of f up to $T - 1$)

6. Lecture 4, slides 7-8 – inconsistency between whether mileage of a new bus is 0 (slide 7) or δ (second formula on slide 8).
7. Lecture 4, slide 23 – different notation for income in real wages; y_t or w_t
8. Lecture 7, slide 27 – no Corollary 4 (same in working paper).

Questions continue:

6. In today's lecture, on slides 21, we use a method that looks like GMM and get a CCP estimator for theta1, which is followed by discussions about its asymptotic covariance matrix. The title of that slides is QMLE, so I'm confused: how is this CCP estimator related to QMLE? And is asymptotic covariance matrix of QMLE also adjusted by a similar fashion?

L5 (Applying CCP to Dynamic Games)

1. Slide 14: how did you derive the conditional value function?

L6 (Representation)

1. Slide 10: what does "DGP is (x_t, d_t)" mean? Aren't (x_t, d_t) the state and data, not a DGP?

L7 (Identification)

1. What does it mean to "know one choice specific payoff for each state for each period"? What is an example of this? If utility is ordinal, why does the value or level of utility matter?

L8 (Finite Dependence)

1. Slide 4: typo on the subscript for f? Should it be $f_{k\ t+tau}$ instead of $f_{k\ tau}$?
2. Slide 18: Should the number of attainable states in period t+1 given either choice i or j be $N_{t+1}(x_t)$ instead of $N_{t+2}(x_t)$?

L9 (CPP Estimators)

1. Slide 18: why does tau in the summation starts with 1 instead of t?