Solutions due by 10.30am Friday 5th March.

1. Show that $\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y \mid X]]$.

   Solution:

$$
\begin{aligned}
\mathbb{E}[\mathbb{E}[Y \mid X]] &= \mathbb{E}\left[\sum_y y\mathbb{P}(Y \mid X)\right] \\
&= \sum_x \left[\sum_y y\mathbb{P}(Y \mid X)\right]\mathbb{P}(X) \\
&= \sum_x \left[\sum_y y\mathbb{P}(Y, X)\right] \\
&= \sum_y y\left[\sum_x \mathbb{P}(Y, X)\right] \\
&= \sum_y y\mathbb{P}(Y) \\
&= \mathbb{E}[Y]
\end{aligned}
$$

2. Suppose $\sqrt{n}(\mathbf{x} - \mu) \xrightarrow{d} N(0, \Sigma)$, where $\mathbf{x}$ is a p-vector and Σ a $p \times p$ covariance matrix. Show that $n(\mathbf{x} - \mu)'\Sigma^{-1}(\mathbf{x} - \mu) \xrightarrow{d} \chi^2(p)$.

   Solution:
   Must assume positive definiteness. By Cholesky decomposition, $\Sigma^{-1} = C'C$. The quadratic form becomes $v'v$ where $v = C\sqrt{n}(\bar{X} - \mu) \sim N(0, C\Sigma C')$. The variance term reduces to $C\Sigma C' = C(C'C)^{-1}C' = C(C^{-1}C'^{-1})C' = (CC^{-1})(C'^{-1}C') = I$. Therefore $v'v \sim \chi^2(p)$.

3. Suppose $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \sigma^2)$. Let $g(x)$ be a differentiable function.

   a) Show that if $g'(\theta) \neq 0$, then $\sqrt{n}(g(\hat{\theta} - g(\theta)) \xrightarrow{d} N(0, g'(\theta)^2\sigma^2)$.

   b) What's the asymptotic distribution of $\bar{X}^2$?

Solution:

a) By the intermediate value theorem, there exists a $c \in [\theta, \hat{\theta}]$ such that

$$\sqrt{n}(g(\hat{\theta}) - (\theta)) = g'(c)\sqrt{n}(\hat{\theta} - \theta)$$

Since $\hat{\theta} \xrightarrow{p} \theta$ and $\theta \leq c \leq \hat{\theta}$, the squeeze theorem tells us that $c \xrightarrow{p} \theta$. Since $\sqrt{n}(\hat{\theta} - \theta)$ is asymtotically normal, Slutsky's theorem tells us that

$$\sqrt{n}(g(\hat{\theta}) - g(\theta)) \xrightarrow{d} N(0, g'(\theta)^2\sigma^2)$$

In order to not have the distribution collapse, we must have $g'(\theta) \neq 0$.

b) If $\mu = \mathbb{E}[X] \neq 0$, set $g(x) = x^2$. $g'(\mu) = 2\mu$. Appealing to the central limit theorem for asymptotic normality and applying the above theorem gives $\sqrt{n}(\bar{X}^2 - \mu^2) \xrightarrow{d} N(0, 4\mu^2\sigma^2)$.

If $\mu = 0$, then $\sqrt{n}\bar{X} \xrightarrow{d} N(0, \sigma^2)$, so $\frac{\sqrt{n}\bar{X}}{\sigma} \xrightarrow{d} N(0, 1)$. Therefore $\left(\frac{\sqrt{n}\bar{X}}{\sigma}\right)^2 = \frac{n\bar{X}^2}{\sigma^2} \xrightarrow{d} \chi^2(1)$, so $n\bar{X}^2 \xrightarrow{d} \sigma^2\chi^2(1)$.

4. Let $X \sim N(\mu, \sigma^2)$. Show that $\mathbb{P}(|X - \mu| > k\sigma) \leq 2e^{-k^2/2}$.

   [Hint: Markov's inequality + MGF]

   Solution:
   Markov's inequality gives tells us that $\mathbb{P}(X > k) \leq \frac{\mathbb{E}[X]}{k}$, if X is a non-negative rv. Let Z be a standard normal deviate, then

   $$\mathbb{P}(Z > k) = \mathbb{P}(e^{tZ} > e^{tk}) \leq \frac{\mathbb{E}[e^{tZ}]}{e^{tk}} = e^{-tk}e^{t^2/2} = e^{t^2/2 - tk}$$

   Since this holds for all t, we can choose a particular t to minimize the RHS. Note that it is minimized when $t = k$ (take derivative, set to zero, check second order conditions). So $\mathbb{P}(Z > k) \leq e^{-k^2/2}$. Note that $\mathbb{P}(|Z| > k) = \mathbb{P}(Z > k) + \mathbb{P}(Z < -k)$. Since Z is symmetric, these probabilities are the same, so

   $$\mathbb{P}(|X - \mu| > \sigma k) = \mathbb{P}(|Z| > k) \leq 2e^{-k^2/2}$$

5. The clinical trial research for the Pfizer vaccine was published in the New England Journal of Medicine: 10.1056/NEJMoa2034577. The trial consists of two populations: the treated

(who received the vaccine) and the placebo (who did not). The study measures the number of COVID infections in each population one week after receiving the second dose.

Let p_T and p_P be the true proportion of people who would became infected with COVID if treated and not treated respectively. Let n_T and n_P be the number of people in each of the treated and placebo populations respectively, with $\hat{p}_T$ and $\hat{p}_P$ the sample proportions observed in the study. The incident risk ratio is defined to be $IRR = p_T/p_P$. The effectiveness of the vaccine is defined to be $1 - IRR$.

The Pfizer study reports: 8 COVID infections in the treated group of size $n_T = 18,198$; 162 infections in the placebo group of size $n_P = 18,325$.

a) What is the point estimate for vaccine effectiveness: $1 - \widehat{IRR}$?

b) What asymptotic distribution does $\hat{p} - p$ have for each population?

c) What asymptotic distribution does $\ln \hat{p}_T - \ln \hat{p}_P$ have?

d) Find a 95% confidence interval for the vaccine effectiveness.

Solution:
a) $\hat{p}_T = \frac{8}{18196}$, $\hat{p}_P = \frac{162}{18325}$. $1 - \widehat{IRR} = 1 - \frac{\frac{8}{18196}}{\frac{162}{18325}} \approx 95.0\%$

b) Note that sample proportions are themselves sample means: $\hat{p} = \frac{\#infections}{n} = \frac{1}{n}\sum_i I_i$ where I_i is a Bernoulli random variable for whether person i became infected. This gives $\mathbb{E}[\hat{p}] = p$ and $\text{Var}(\hat{p}) = \frac{p(1-p)}{n}$. Appealing to the central limit theorem gives

$$\sqrt{n_T}(\hat{p}_T - p_T) \xrightarrow{d} N(0, p_T(1 - p_T)) \qquad \sqrt{n_P}(\hat{p}_P - p_P) \xrightarrow{d} N(0, p_P(1 - p_P))$$

c) Let $g(x) = \ln x$, then $g'(x) = \frac{1}{x}$. We proved the Delta method above, yielding

$$\sqrt{n_T}(\ln \hat{p}_T - \ln p_T) \xrightarrow{d} N(0, \frac{1 - p_T}{p_T}) \qquad \sqrt{n_P}(\ln \hat{p}_P - \ln p_P) \xrightarrow{d} N(0, \frac{1 - p_P}{p_P})$$

Moving $\sqrt{n}$s to the other side and differencing gives

$$(\ln \hat{p}_T - \ln p_T) - (\ln \hat{p}_P - \ln p_P) \xrightarrow{d} N(0, \frac{1 - p_T}{p_T n_T} + \frac{1 - p_P}{p_P n_P})$$

Or

$$\ln \hat{p}_T - \ln \hat{p}_P \overset{d}{\to} N(p_T - p_P, \frac{1 - p_T}{p_T n_T} + \frac{1 - p_P}{p_P n_P})$$

c) Note that $1 - \exp \ln \hat{p}_T - \ln \hat{p}_P = 1 - \widehat{IRR}$, which is our point estimator for effectiveness. So to make a CI for effectiveness, let's make a CI for the log difference, exponentiate it and subtract that from one. The log difference CI is

$$\left((\ln \hat{p}_T - \ln \hat{p}_P) - z_{1-\alpha/2} \left(\frac{1 - p_T}{p_T n_T} + \frac{1 - p_P}{p_P n_P} \right), (\ln \hat{p}_T - \ln \hat{p}_P) + z_{\alpha/2} \left(\frac{1 - p_T}{p_T n_T} + \frac{1 - p_P}{p_P n_P} \right) \right)$$

Estimate the standard deviation with our sample proportions: $\sqrt{\frac{1-\hat{p}_T}{\hat{p}_T n_T} + \frac{1-\hat{p}_P}{\hat{p}_P n_P}} \approx 0.362,$. The log difference is approx -3.00, and at the 95% confidence level, $z_{\alpha/2} = 1.96$. This yields a CI of $(-3.71, -2.29)$ for the log difference. Apply our $1 - \exp \cdot$ transform to each end point and we get $(.898, 0.975)$. So we can be 95% confident that vaccine effectiveness is between 89.8% and 97.5%. Note that the CI found by the authors is (90.3%, 97.6%), they used a slightly different type of confidence interval (not strictly a confidence interval, but a Bayesian credible interval).

6. Suppose an agent needs to know a parameter θ, but cannot observe it directly. Suppose they are given two noisy signals:

$$x = \theta + \varepsilon$$
$$y = \theta + \nu$$

where $\theta \sim N(0, \sigma_\theta^2)$, $\varepsilon \sim N(0, \sigma_\varepsilon^2)$, $\nu \sim N(0, \sigma_\nu^2)$, all pairwise independent. The agent estimates θ with a linear combination of the signals. What weights should they choose such that their estimator is minimum mean square error?

Solution:
Look for an estimator of the form $\hat{\theta} = c_1 x + c_2 y$. Apply the orthogonality principle. We need:

$$\mathbb{E}[(\hat{\theta} - \theta) \cdot x] = 0$$
$$\mathbb{E}[(\hat{\theta} - \theta) \cdot y] = 0$$

This gives

$$c_1(\sigma_\theta^2 + \sigma_\varepsilon^2) + c_2\sigma_\theta^2 = \sigma_\theta^2$$
$$c_1\sigma_\theta^2 + c_2(\sigma_\theta^2 + \sigma_\nu^2) = \sigma_\theta^2$$

Or

$$\begin{pmatrix} \sigma_\theta^2 + \sigma_\varepsilon^2 & \sigma_\theta^2 \\ \sigma_\theta^2 & \sigma_\theta^2 + \sigma_\nu^2 \end{pmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \sigma_\theta^2 \\ \sigma_\theta^2 \end{bmatrix}$$

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \frac{1}{\sigma_\theta^2\sigma_\nu^2 + \sigma_\theta^2\sigma_\varepsilon^2 + \sigma_\varepsilon^2\sigma_\nu^2} \begin{pmatrix} \sigma_\theta^2 + \sigma_\nu^2 & -\sigma_\theta^2 \\ -\sigma_\theta^2 & \sigma_\theta^2 + \sigma_\varepsilon^2 \end{pmatrix} \begin{bmatrix} \sigma_\theta^2 \\ \sigma_\theta^2 \end{bmatrix}$$

Which gives

$$c_1 = \frac{\frac{1}{\sigma_\varepsilon^2}}{\frac{1}{\sigma_\varepsilon^2} + \frac{1}{\sigma_\nu^2} + \frac{1}{\sigma_\theta^2}} \qquad c_2 = \frac{\frac{1}{\sigma_\nu^2}}{\frac{1}{\sigma_\varepsilon^2} + \frac{1}{\sigma_\nu^2} + \frac{1}{\sigma_\theta^2}}$$

7. Consider the Solow growth model described by

$$k_{t+1} = sk_t^\alpha + (1 - \delta)k_t \tag{1}$$
$$c_t = (1 - s)k_t^\alpha \tag{2}$$

where $0 \leq \alpha < 1$, $s \in [0, 1]$ is a constant savings rate, and $\delta \in [0, 1]$ is the capital depreciation rate.

a) Solve for the unique steady state levels of capital and consumption.

b) Find the savings rate that maximizes consumption in the steady state.

c) Log-linearize the two equations.

d) Describe the approximate solution as a system of first order difference equations.

Solution:
a) $\bar{k} = s\bar{k}^\alpha + (1 - \delta)\bar{k} \implies \delta\bar{k} = s\bar{k}^\alpha \implies \bar{k} = (\frac{\delta}{s})^{\frac{1}{\alpha-1}} \implies \bar{k} = (\frac{s}{\delta})^{\frac{1}{1-\alpha}}$. Then $\bar{c} = (1 - s)\bar{k}^\alpha$.

b) $\bar{(c)} = \bar{k}^\alpha - s\bar{k}^\alpha$. From the capital flow equation we know that in the steady state $s\bar{k}^\alpha = \delta\bar{k}$. Let's sub that in to geh $\bar{c} = \bar{k}^\alpha - \delta\bar{k}$. Differentiating wrt the savings rate gives

$$\frac{d\bar{c}}{ds} = [\alpha\bar{k}^{\alpha-1} - \delta]\frac{d\bar{k}}{ds} = 0$$

Note that $\frac{d\bar{k}}{ds} = 0$ only if $s = 0$. So at the optimum we must have $\alpha\bar{k}^{\alpha-1} = \delta$. So $\alpha\left[\left(\frac{s}{\delta}\right)^{\frac{1}{1-\alpha}}\right]^{\alpha-1} = \delta \implies s = \alpha$.

c) The capital flow equation becomes: $\bar{k}\hat{k}_{t+1} = \left[\alpha s\bar{k}^{\alpha-1} + 1 - \delta\right]\bar{k}\hat{k}_t$. The consumption equation becomes $\bar{c}\hat{c}_t = \alpha(1-s)\bar{k}^{\alpha-1}\bar{k}\hat{k}_t$.

d)
$$\begin{pmatrix} 1 & -\alpha(1-s)\bar{k}^{\alpha-1}\bar{k} \\ 0 & 1 \end{pmatrix}\begin{bmatrix} c_{t+1} \\ k_{t+1} \end{bmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \alpha s\bar{k}^{\alpha-1} + 1 - \delta \end{pmatrix}\begin{bmatrix} c_t \\ k_t \end{bmatrix}$$

8. Log-linearize the following equations assuming a steady state exists:

   a) $c_t + k_{t+1} = F(k_t) + (1-\delta)k_t$

   b) $U'(c_t) = \beta U'(c_{t+1})[F'(k_{t+1}) + 1 - \delta]$

   Solution:

   a) $\bar{c}\hat{c}_t + \bar{k}\hat{k}_{t+1} = \left[F'(\bar{k}_t) + 1 - \delta\right]\bar{k}\hat{k}_t$. Noting that in steady state $1 = \beta[F'(\bar{k}) + 1 - \delta]$, this reduces to $\bar{c}\hat{c}_t + \bar{k}\hat{k}_{t+1} = \beta^{-1}\bar{k}\hat{k}_t$

   b) $U''(\bar{c})\bar{c}\hat{c}_t = \beta U''(\bar{c})[F'(\bar{k}) + 1 - \delta]\bar{c}\hat{c}_{t+1} + \beta U'(\bar{c})F''(\bar{k})\bar{k}\hat{k}_{t+1}$.

   Divide through by $U'(\bar{c})$ and note that in steady state $1 = \beta[F'(\bar{k}) + 1 - \delta]$.

   $\sigma(\bar{c})c_t = \sigma(\bar{c})\hat{c}_{t+1} + \beta F''(\bar{k})\bar{k}\hat{k}_{t+1}$, where $\sigma(\bar{c}) = \frac{U''(\bar{c})\bar{c}}{U'(\bar{c})}$, is the Arrow-Pratt measure of risk aversion.