10.1　随机游动的非平稳性表现在哪里？为什么带漂移项的随机游动隐含着时间趋势？

10.2　$\{w_t\}_{t=1}^T$ 为标准随机游动 $w_t = w_{t-1} + \varepsilon_t$，$\varepsilon_t$ 为白噪声。由于数学期望为 0，$\{w_t\}_{t=1}^T$ 的自协方差为 $\mathrm{cov}(w_t, w_s) = Ew_t w_s$。计算 $\{w_t\}_{t=1}^T$ 的自协方差函数和自相关函数。你有什么发现？这说明了什么？

10.3　什么是伪回归？为什么会出现伪回归？

10.4　设 $y_t = c + \delta t + \rho y_{t-1} + \varepsilon_t$，$0 < \rho < 1$ 为趋势平稳时间序列。求出 y_t 的自协方差函数和自相关函数。你有什么发现？为什么说趋势平稳时间序列本质上是平稳序列？

10.5　单位根 ADF 检验的检验模型有几种？分别适用于什么数据？选用的检验模型和数据不一致会出现什么结果？

10.6　如何确定 ADF 检验中的滞后阶数？

10.7　如果没有实际情况信息可供参考，单位根检验应该遵循怎样的流程？为什么？如何确定模型是否包含常数项和时间趋势项？

10.8　为什么差分可以减轻序列不平稳的程度？什么是单整？什么是 ARIMA 模型？

10.9　什么是协整？叙述 E-G 两步法检验协整的原理和步骤。协整回归中需要包含常数项和时间趋势项吗？为什么？

10.10 什么是误差修正模型 ECM？与 ARIMA 模型相比，ECM 模型有什么优点？如何建立 ECM 模型？

10.11 从《中国统计年鉴》或者中经网获取中国人口年度时间序列数据。对人口序列进行单位根检验，建立和估计其 ARIMA 模型。

10.12 对 20012 年 1 月 1 日至 2013 年 1 月 30 招商银行（600036）日收盘价的对数形成的时间序列进行单位根检验，与同一时期的沪综指收盘对数指数进行协整分析，从中你能得出什么结论？

10.13 对 2012 年 1 月 1 日至 2013 年 1 月 30 的沪深 300（000300）日收盘对数指数（指数的自然对数）时间序列进行平稳性检验，建立和估计其 ARIMA 模型。

10.14*当单整时间序列 x_t 和 y_t 存在协整关系时，证明误差修正模型（10.19 自回归分布滞后模型

$$y_t = \beta_0 + \beta_1 y_{t-1} + \alpha_0 x_t + \alpha_1 x_{t-1} + \varepsilon_t \tag{10.21}$$

等价。将（10.21）表示成 ECM 形式。

◆ 参考答案

1. 随机游动的非平稳性体现在序列的方差随时间变化，主要原因在于其自回归系数等于 1，$y_t = y_{t-1} + \varepsilon_t$，误差项 ε_s 随时间推移对 y_{s+k} 的影响没有丝毫衰减，y_t 等于到 t 为止的误差之和，其方差是各期误差项方差之和，是 t 的倍数，随 t 的增加而增加。带漂移项的随机游动，其漂移项的累计与标准随机游动方差的累计一样，是 t 的倍数，形成序列的时间趋势。

2. 设 $t > s$，则 $w_t = w_{t-1} + \varepsilon_t = \cdots = w_s + \varepsilon_{s+1} + \cdots + \varepsilon_{s+(t-s)}$，由于 $\{\varepsilon_t\}_{t=1}^T$ 为白噪声，因此 $\{w_t\}_{t=1}^T$ 的协方差为 $\mathrm{cov}(w_t, w_s) = \mathrm{E}w_t w_s = \mathrm{E}w_s(w_s + \varepsilon_{s+1} + \cdots + \varepsilon_{s+(t-s)}) = \mathrm{E}w_s^2 = s\sigma_\varepsilon^2$，类似的，当 $s > t$ 时，$\mathrm{cov}(w_t, w_s) = \mathrm{E}w_t w_s = t\sigma_\varepsilon^2$。因此 $\mathrm{cov}(w_t, w_s) = \min(t, s)\sigma_\varepsilon^2$。自协方差函数为

$$\rho(t,s)=\rho(w_t,w_s)=\frac{\mathrm{Cov}(w_t,w_s)}{\sqrt{\mathrm{Var}(w_t)\mathrm{Var}(w_s)}}$$

$$=\frac{\min(t,s)\sigma_\varepsilon^2}{\sqrt{t\sigma_\varepsilon^2}\sqrt{s\sigma_\varepsilon^2}}=\frac{\min(t,s)}{\sqrt{st}}=\begin{cases}\sqrt{s/t} & ,t>s\\ \sqrt{t/s} & ,s>t\end{cases}=\begin{cases}\sqrt{s/(s+k)} & ,t>s\\ \sqrt{t/(t+k)} & ,s>t\end{cases}$$

其中 $k=\max(t,s)-\min(t,s)$。由此可以看出，单位根时间序列的自相关函数随时间点 (t,s) 的变化而变化，体现出非平稳时间序列的特点。从自相关函数中看出，w_t 和 w_s 的相关函数随时间间隔 $k=\max(t,s)-\min(t,s)$ 的增加而递减，并且当 $k\to\infty$ 时，$\rho(s,t)\to 0$。

3. 伪回归是指没有线性关系的时间序列变量，回归系数的 t-检验显示回归系数显著不为 0，得出变量间具有显著线性关系的错误结论。造成伪回归的原因，是回归变量的非平稳性，参数的 OLS 估计不再渐近服从正态分布，以此构造的 t-统计量不服从 t-分布，按 t-分布临界值得出的检验结果错误。
4. 由于 δt 为非随机项，不影响 y_t 的方差，也不影响 y_t 和 y_s 的协方差。y_t 的自协方差函数和自相关函数与平稳时间序列 AR(1) $y_t=c+\rho y_{t-1}+\varepsilon_t$，$0<\rho<1$ 相同，自协方差函数为 $\gamma_k=\rho_1^k\sigma_\varepsilon^2/(1-\rho_1^2), k=1,2,\ldots$，自相关函数为 $\rho(k)=\phi_1^k, k=1,2,\ldots$。自相关函数和自协方差函数都与时间 t 无关，只与时间间隔 k 有关，表现出平稳时间序列的特征。唯一不满足平稳性的是数学期望随 t 的变化而变化，而这种变化是由非随机的时间趋势 δt 引起的，从本质上说，趋势平稳是平稳时间序列。
5. ADF 检验的检验模型有三种。数据的平均值为 0（数据围绕 x 轴上下波动），且没有明显的时间趋势，采用既没有常数项又没有时间趋势项的模型进行检验。数据的平均值不为 0，且没有明显时间趋势，采用带常数项不带时间趋势项的模型进行检验。数据有明显的时间趋势，需要确定序列是趋势平稳还是带常数项非平稳时间序列时，采用既带常数项又带时间趋势项的模型进行检验。如果选用的检验模型和数据不一致，检验采用的统计量是错误的，得出的临界值及检验结果也是错误的。
6. 采用信息准则确定 ADF 检验模型中 Δy_t 的滞后阶数，常用的有 AIC 准则和 SC 准则，SC 准则得出的滞后阶数较小。默认采用 SC 准则。
7. 应该采用从既包含常数项又包含时间趋势项的检验模型开始检验，逐步确定检验模型是否应该包含常数项和时间趋势项，并最终得出检验结论。因为，只有确定了合适的检验模型，检验结果才是正确的。采用参数约束检验中的 T_r 统计量来确定检验模型是否包含常数项和时间趋势项，但此时 T_r 不再服从 F 分布，需要查特殊的分布临界值表确定检验的临界值。
8. 序列的不平稳性是由趋势项造成的，不管是确定性时间趋势，还是随机游动形成的随机趋势，确定性时间趋势前后两项相减后成为常数 $\delta(t+1)-\delta t=\delta$，随机游动前后两项相减后成为平稳序列 $y_t-y_{t-1}=\varepsilon_t$。单整是指非平稳时间序列经多次差分后平稳，差分次数就是单证阶数。ARIMA 模型是采用差分平稳后的序列建立的自回归模型。
9. 协整是指同阶单整的两个时间序列线性组合后成为平稳时间序列。E-G 两步法的第一步对两个时间序列变量进行 OLS 回归，第二部对回归残差进行平稳性检验。如果残差为平稳序列，则两个变量存在协整关系，回归模型的估计结果就是协整组合。协整回归中需要包含常数项和时间趋势项，这样得出的回归残差均值为 0，且不再有时间趋势。在第二部对残差的单位根检验中，只需要选择既无常数项又无时间趋势项的检验模型即可。

10. ECM 模型是将自变量差分平稳后的变量和协整组合形成的（平稳）变量作为自变量，以因变量差分平稳后的变量作为因变量的回归模型。由于增加量误差修正项，ECM 模型比 ARIMA 模型更为灵活，不仅刻画了两个变量的短期动态关系，也刻画了长期动态关系变化对短期变化的调整作用。要建立 ECM 模型，需首先确定变量的单整阶数为 1，然后检验变量间存在协整关系，并估计出协整组合，并把协整组合和自变量差分一起作为自变量，以因变量差分作为因变量建立回归模型，进行 OLS 估计得出 ECM 模型。

11.、12.、13. 略

14* 将 $\Delta y_t = (\gamma_0 - \gamma_2\alpha) + \gamma_1 \Delta x_t + \gamma_2 (y_{t-1} - \beta x_{t-1}) + v_t$ 中的差分项写开，并整理得出

$$\begin{aligned} y_t &= (\gamma_0 - \gamma_2\alpha) + y_{t-1} - \gamma_2 y_{t-1} + \gamma_1 x_t - \gamma_1 x_{t-1} + \gamma_2 y_{t-1} - \gamma_2 \beta x_{t-1}) + v_t \\ &= (\gamma_0 - \gamma_2\alpha) + (1-\gamma_2) y_{t-1} + \gamma_1 x_t - (\gamma_1 + \gamma_2 \beta) x_{t-1} + v_t \\ &= \beta_0 + \beta_1 y_{t-1} + \alpha_0 x_t + \alpha_1 x_{t-1} + \varepsilon_t \end{aligned}$$

，即（10.21）得出

将（10.21）做如下变形

$$\begin{aligned} y_t - y_{t-1} &= \beta_0 + \beta_1 y_{t-1} + \alpha_0 x_t + \alpha_1 x_{t-1} + \varepsilon_t \\ &= \beta_0 + (\beta_1 - 1) y_{t-1} + \alpha_0 (x_t - x_{t-1}) + (\alpha_1 + \alpha_0) x_{t-1} + \varepsilon_t \\ &= \beta_0 + \alpha_0 \Delta x_t + (\beta_1 - 1)\left(y_{t-1} - \frac{\alpha_1 + \alpha_0}{1 - \beta_1} \right) x_{t-1} + \varepsilon_t \end{aligned}$$

，成为 ECM 形式。