Time Series: midterm, spring 25'

1. **Условие:**

   При каких p и q указанный процесс будет являться белым шумом?

   $$X_t = A_t \cdot (-1)^{S_t}$$

   где:

   - A_t – i.i.d. $Bern(p)$
   - $S_t = \sum\limits_{k=1}^{t-1} B_k$, где B_k - i.i.d. $Bern(q)$
   - B_k и A_t независимы

   Решение:

   (1) Если X_t – белый шум, то по определению $\forall t : \mathbb{E}X_t = 0$. Найдём значения p и q, удовлетворяющие этому условию. $\mathbb{E}X_t = \mathbb{E}A_t \cdot (-1)^{S_t} = \mathbb{E}A_t \cdot \mathbb{E}(-1)^{S_t} = \mathbb{E}A_t \cdot \mathbb{E}(-1)^{\sum\limits_{k=1}^{t-1} B_k} = \mathbb{E}A_t \cdot \prod\limits_{k=1}^{t-1} \mathbb{E}(-1)^{B_k}$, где цепочка равенств следует из того, что $\mathbb{E}(\xi\eta) = \mathbb{E}\xi \cdot \mathbb{E}\eta$ для любых независимых случайных величин ξ и η.

   $$\left.\begin{array}{l} \mathbb{E}A_t = 0 \cdot (1-p) + 1 \cdot p = p1 \\ \mathbb{E}(-1)^{B_k} = 1 \cdot (1-q) + (-1) \cdot q = 1 - 2q \end{array}\right\} \Rightarrow \mathbb{E}X_t = p \cdot (1-2q)^{t-1} = 0 \Rightarrow (p = 0) \vee (q = \tfrac{1}{2})$$

   (2) При $p = 0$: $A_t \sim Bern(0) \Rightarrow A_t \equiv 0 \Rightarrow X_t \equiv 0$. В таком случае, X_t – белый шум, так как мат. ожидание константы равно ей самой и равно 0, а ковариация константы с любой другой случайной величиной равно 0, то есть $\forall t, s : cov(X_t, X_s) = 0$, в том числе $\mathbb{V}arX_t = 0$. Значит, при $p = 0$ величина q может быть произвольна. Таким образом, $p = 0, q \in [0, 1]$ – один из ответов.

   (3) При $p > 0$ необходимо $q = \frac{1}{2}$, оценим дисперсию и ковариацию.
   $\mathbb{V}arX_t = \mathbb{E}X_t^2 - (\mathbb{E}X_t)^2 = \mathbb{E}A_t^2(-1)^{2S_t} = \mathbb{E}A_t^2 = 0 \cdot (1-p) + 1 \cdot p = p$

   $cov(X_t, X_s) = \{\text{t>s}\} = \mathbb{E}(X_t - \mathbb{E}X_t)(X_s - \mathbb{E}X_s) = \mathbb{E}X_tX_s = \mathbb{E}A_tA_s \cdot (-1)^{S_t + S_s} = \mathbb{E}A_t \cdot \mathbb{E}A_s \cdot \mathbb{E}(-1)^{\sum\limits_{i=1}^{t-1} B_i + \sum\limits_{j=1}^{s-1} B_j} = p^2 \cdot \mathbb{E}(-1)^{2\sum\limits_{i=1}^{s-1} B_i + \sum\limits_{j=s}^{t-1} B_j} = p^2 \mathbb{E}(-1)^{\sum\limits_{j=s}^{t-1} B_j} = p^2 \cdot \prod\limits_{j=s}^{t-1} \mathbb{E}(-1)^{B_j} = p^2 \cdot (1-2q)^{t-s} = 0$ при $q = \frac{1}{2}, t > s$.
   Значит, $p \in (0, 1], q = \frac{1}{2}$ удовлетворяет условиям белого шума.

   Замечание: с учётом дословного следования текущему условию вариант $p > 0$ не может быть решением, так как $X_1 = A_1 \cdot (-1)^{\sum\limits_{k=1}^{0} B_k} = A_1 \cdot (-1)^0 = A_1$, откуда $\mathbb{E}A_1 = p > 0$; однако это небольшая некорректность и правильнее полагать $S_t = \sum\limits_{k=1}^{t} B_k$.

2. **Условие:**

   Пересвет Матрёшкин считает, что его данные могут быть описаны следующим уравнением:

   $$y_t - 0.5y_{t-1} + 0.06y_{t-2} = 10 + \varepsilon_t - 0.9\varepsilon_{t-1} + 0.26\varepsilon_{t-2} - 0.024\varepsilon_{t-3}$$

   (a) Проверьте, сколько у этого уравнения нестационарных, стационарных и стабильных решений?

   (b) Классифицируйте процесс как ARMA. Укажите p и q.

   (c) Найти $ACF(k), PACF(k)$ в явном виде как функцию от k.

(d) Найдите предел $\mathbb{E}(y_{T+h}|\mathcal{F}_T)$ при $h \to \infty$.

(e) Выразите ε_t через предыдущие лаги y_t или докажите, что это невозможно.

Решение:

(a) Перепишем уравнение в виде через лаговые полиномы от $y_t - \mu$ и ε_t.

$$(y_t - \mu) - 0.5(y_{t-1} - \mu) + 0.06(y_{t-2} - \mu) = \varepsilon_t - 0.9\varepsilon_{t-1} + 0.26\varepsilon_{t-2} - 0.024\varepsilon_{t-3} \Rightarrow$$

$$-\mu + 0.5\mu - 0.06\mu = -10 \Rightarrow -0.56\mu = -10, \mu = \frac{1000}{56} = \frac{125}{7} \approx 17.86$$

Тогда исходное уравнение переписывается в виде:

$$(1 - 0.5L + 0.06L^2)(y_t - \mu) = (1 - 0.9L + 0.26L^2 - 0.024L^3)\varepsilon_t \Leftrightarrow$$

$$(1 - 0.2L)(1 - 0.3L)(y_t - \mu) = (1 - 0.2L)(1 - 0.3L)(1 - 0.4L)\varepsilon_t$$

После сокращения скобок получаем: $y_t - \mu = (1 - 0.4L)\varepsilon_t = \varepsilon_t - 0.4\varepsilon_{t-1}$. Полученный процесс является $MA(1)$-процессом он стационарен. Нестационарных решений у него бесконечно много; так как у его лагового многочлена нет корней, равных или меньших единицы по модулю, то его стабильное решение будет стационарным.

Таким образом, имеем бесконечно много нестационарных решений, 1 стационарное и 1 стабильное.

(b) $y_t = \mu + \varepsilon_t - 0.4\varepsilon_{t-1} \Rightarrow y_t$ – AR(0) и MA(1) процесс, то есть ARMA(0, 1).

Процесс нельзя определить как ARMA(2, 3), используя начальное соотношение, так как в определении ARMA-процесса через равенство вида $A(L)(y_t - \mu) = B(L)\varepsilon_t$ полиномы $A(L)$ и $B(L)$ полагаются несократимыми, что не верно для начального уравнения.

(c) $ACF(k) = \rho_k$, положим $\gamma_k = cov(y_t, y_{t-k})$, тогда $\rho_k = \frac{\gamma_k}{\gamma_0}$.

- $\gamma_0 = cov(\mu+\varepsilon_t-0.4\varepsilon_{t-1}, \mu+\varepsilon_t-0.4\varepsilon_{t-1}) = cov(\varepsilon_t-0.4\varepsilon_{t-1}, \varepsilon_t-0.4\varepsilon_{t-1}) = cov(\varepsilon_t, \varepsilon_t)+0.16cov(\varepsilon_{t-1}, \varepsilon_{t-1}) = 1.16\sigma^2$
- $\gamma_1 = cov(\mu+\varepsilon_t-0.4\varepsilon_{t-1}, \mu+\varepsilon_{t-1}-0.4\varepsilon_{t-2}) = cov(\varepsilon_t-0.4\varepsilon_{t-1}, \varepsilon_{t-1}-0.4\varepsilon_{t-2}) = -0.4cov(\varepsilon_{t-1}, \varepsilon_{t-1}) = -0.4\sigma^2$
- $\gamma_k = 0$ при $k \geq 2$, так как y_t – MA(1)-процесс.
- $\rho_0 = 1, \rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{-0.4\sigma^2}{1.16\sigma^2} = -\frac{10}{29} \approx -0.34$, $\rho_k = \frac{\gamma_k}{\gamma_0} = 0, k \geq 2$.

$PACF(k) = \varphi_{kk}$

- $\varphi_{11} = \rho_1 = -\frac{10}{29} \approx -0.34$
- Для φ_{kk} найдём общую формулу через теорему Юла-Волкера.
  $y_t = m + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots + \alpha_k y_{t-k} + \eta_t \Rightarrow \eta_t = y_t - m - \alpha_1 y_{t-1} - \alpha_2 y_{t-2} - \ldots - \alpha_k y_{t-k}$

$$\begin{cases} cov(\eta_t, y_{t-1}) = 0 \\ \ldots \\ cov(\eta_t, y_{t-k}) = 0 \end{cases} \Leftrightarrow \begin{cases} cov(y_t - m - \alpha_1 y_{t-1} - \alpha_2 y_{t-2} - \ldots - \alpha_k y_{t-k}, y_{t-1}) = 0 \\ \ldots \\ cov(y_t - m - \alpha_1 y_{t-1} - \alpha_2 y_{t-2} - \ldots - \alpha_k y_{t-k}, y_{t-k}) = 0 \end{cases}$$

$$\begin{cases} \gamma_1 - \alpha_1\gamma_0 - \alpha_2\gamma_1 - \ldots - \alpha k\gamma_{k-1} = 0 \\ \ldots \\ \gamma_k - \alpha_1\gamma_{k-1} - \alpha_2\gamma_{k-2} - \ldots - \alpha_k\gamma_0 = 0 \end{cases}$$

  Так как $\gamma_k = 0, k \geq 2$, то это уравнение можно переписать в виде:

$$\begin{pmatrix} \gamma_0 & \gamma_1 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\ \gamma_1 & \gamma_0 & \gamma_1 & 0 & \ldots & 0 & 0 & 0 & 0 \\ 0 & \gamma_1 & \gamma_0 & \gamma_1 & \ldots & 0 & 0 & 0 & 0 \\ \ldots \\ \ldots \\ 0 & 0 & 0 & 0 & \ldots & 0 & \gamma_1 & \gamma_0 & \gamma_1 \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 & \gamma_1 & \gamma_0 \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \\ \ldots \\ \alpha_{k-1} \\ \alpha_k \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ 0 \\ \ldots \\ \ldots \\ 0 \\ 0 \end{pmatrix}$$

- Тогда найти α_k можно по формуле Крамера как

$$\alpha_k = \varphi_{kk} = \frac{det\begin{pmatrix} \gamma_0 & \gamma_1 & 0 & 0 & \dots & 0 & 0 & 0 & \gamma_1 \\ \gamma_1 & \gamma_0 & \gamma_1 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & \gamma_1 & \gamma_0 & \gamma_1 & \dots & 0 & 0 & 0 & 0 \\ \dots \\ \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & \gamma_1 & \gamma_0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & \gamma_1 & 0 \end{pmatrix}}{det\begin{pmatrix} \gamma_0 & \gamma_1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ \gamma_1 & \gamma_0 & \gamma_1 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & \gamma_1 & \gamma_0 & \gamma_1 & \dots & 0 & 0 & 0 & 0 \\ \dots \\ \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & \gamma_1 & \gamma_0 & \gamma_1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & \gamma_1 & \gamma_0 \end{pmatrix}}$$

- Вычисляя определители матриц как реккуренты, получаем общую формулу

$$\varphi_{kk} = \frac{-(-\theta)^k}{1+\theta^2+\ldots+\theta^{2k}}$$

где θ – параметр из MA(1) процесса $y_t - \mu = \varepsilon_t + \theta\varepsilon_{t-1}$, откуда

$$\varphi_{kk} = \frac{-(-(-0.4))^k}{\frac{1-(-0.4)^{2k+2}}{1-(-0.4)^2}} = \frac{-0.4^k \cdot 0.84}{1-0.4^{2k+2}}$$

На контрольной было достаточно найти PACF для первых 2 шагов; рассуждения полностью повторяют вывод общей формулы и опираются на теорему Юла-Волкера.

`https://stats.stackexchange.com/questions/140371/pacf-for-ma1-process`

(d) Заметим, что $\mathcal{F}_T$ зависит от $\varepsilon_t, t \le T$, откуда $\lim\limits_{h\to\infty} \mathbb{E}(y_{T+h}|\mathcal{F}_T) = \lim\limits_{h\to\infty} \mathbb{E}(\mu + \varepsilon_{T+h} - 0.4\varepsilon_{T+h-1}|\mathcal{F}_T) = \mu + \lim\limits_{h\to\infty} \mathbb{E}(\varepsilon_{T+h}|\mathcal{F}_T) - 0.4 \lim\limits_{h\to\infty} \mathbb{E}(\varepsilon_{T+h-1}|\mathcal{F}_T) = \mu$

(e) $y_t - \mu = (1-0.4L)\varepsilon_t \Rightarrow \varepsilon_t = (1-0.4L)^{-1}(y_t - \mu) = (1-0.4L)^{-1}y_t - \frac{\mu}{1-0.4L} = \sum\limits_{k=0}^{\infty}(0.4L)^k y_t - \frac{5}{3}\mu = -\frac{5}{3}\mu + \sum\limits_{k=0}^{\infty} 0.4^k y_{t-k}$

3. **Условие:**

Рассмотрим модель ETS(AAdN):

$$\begin{cases} u_t \sim \mathcal{N}(0;20) \\ b_t = 0.8b_{t-1} + 0.3u_t \\ \ell_t = \ell_{t-1} + 0.8b_{t-1} + 0.2u_t \\ y_t = \ell_{t-1} + 0.8b_{t-1} + u_t \end{cases}$$

где $\ell_{100} = 20, b_{100} = 1$.

(a) Найдите 95% доверительный интервал y_{102}.

(b) Приблизительно вычислите прогноз для y_{2025}.

Решение:

(a) $b_{101} = 0.8b_{100} + 0.3u_{101} \Rightarrow \mathbb{E}(b_{101}|\mathcal{F}_{100}) = 0.8 + 0.3u_{101}$
$\ell_{101} = \ell_{100} + 0.8b_{100} + 0.2u_{101} \Rightarrow \mathbb{E}(\ell_{101}|\mathcal{F}_{100}) = 20 + 0.8 + 0.2u_{101} = 20.8 + 0.2u_{101}$
$y_{102} = \ell_{101} + 0.8b_{101} + u_{102} = (\ell_{100} + 0.8b_{100} + 0.2u_{101}) + 0.8(0.8b_{100} + 0.3u_{101}) + u_{102} = \ell_{100} + 1.44b_{100} + 0.44u_{101} + u_{102} \Rightarrow \mathbb{E}(y_{102}|\mathcal{F}_{100}) = 20 + 1.44 \cdot 1 + 0.44u_{101} + u_{102} = 21.44 + 0.44u_{101} + u_{102}$ Заметим, что u_{101} и u_{102} – некоррелированные, откуда $0.44u_{101} + u_{102} \sim \mathcal{N}(0; 0.44^2 \cdot 20 + 20) = \mathcal{N}(0; 23.872)$.
$\mathbb{E}(y_{102}|\mathcal{F}_100) \sim \mathcal{N}(21.44; 23.872) \Rightarrow$; с вероятностью 0.95 :

$$y_{102} \in (21.44 - 1.96 \cdot \sqrt{23.872}; 21.44 + 1.96 \cdot \sqrt{23.872}) \approx (11.86; 31.02)$$

(b) В качестве прогноза модели для шага y_{2025} будем использовать $\mathbb{E}(y_{2025}|\mathcal{F}_100)$, заметим, что на каждом шаге к y_t прибавляются слагаемых u_t, а через ℓ_{t-1} с b_{t-1} прибавляются и u_{t-1}. Однако, они будут присутствовать в равенстве в первых степенях и в силу $\mathbb{E}u_t = 0$ сократятся. Значит, для поиска математического ожидания достаточно выразить y_{2025} через ℓ_{100} и b_{100}, опустив u_t.

Тогда $y_t = \ell_{t-1} + 0.8b_{t-1} = (\ell_{t-2} + 0.8b_{t-2}) + 0.8 \cdot 0.8b_{t-2} = \ldots = \ell_{t-k} + b_{t-k}\sum\limits_{i=1}^{k} 0.8^i = \ell_{t-k} + b_{t-k}\frac{0.8-0.8^{k+1}}{1-0.8}$.
$y_{2025} = \ell_{2025-1925} + b_{2025-1925} \cdot \frac{0.8-0.8^{1926}}{1-0.8} \approx \ell_{100} + b_{100} \cdot \frac{0.8}{0.2} = \ell_{100} + 4b_{100} = 24$

4. **Условие:**
У Лукоморья дуб зелёный. Златая цепь на дубе том. Пусть H– размах кроны дуба в кошачьих шагах. Каждые H шагов кот учёный меняет направление своего движения и тема его повествования меняется. ρ – параметр креативности кота. $|\rho| < 1$. Тогда историю можно описать следующим процессом:

$$x_t = \begin{cases} \rho x_{t-1} + \varepsilon_t, & mod(t/H) \neq 0 \\ \varepsilon_t, & mod(t/H) = 0 \end{cases}$$

При каком условии на H процесс будет стационарным? ε_t – белый шум.

Решение:

(1) Заметим, что при $t = pH + r, 0 \leq r < H$ верно, что $x_t = \rho x_{t-1} + \varepsilon_t = \rho(\rho x_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \sum\limits_{k=0}^{r} \rho^k \varepsilon_{t-k}$. Таким образом, в силу некоррелированности белого шума: $\mathbb{V}ar x_t = \mathbb{V}ar(\sum\limits_{k=0}^{r} \rho^k \varepsilon_{t-k}) = \sum\limits_{k=0}^{r} \rho^{2k} \mathbb{V}ar(\varepsilon_{t-k}) = \sum\limits_{k=0}^{r} \rho^{2k}\sigma^2$.

(2) При $\rho = 0$: $\forall t : x_t = \varepsilon_t$, откуда x_t – стационарный по свойствам белого шума, то есть H – любой.

(3) При $\rho \neq 0$: $\mathbb{V}ar(x_{kH}) = \mathbb{V}ar(\varepsilon_{kH}) = \sigma^2$ и $\mathbb{V}ar(x_{kH+1}) = \mathbb{V}ar(\rho\varepsilon_{kH} + \varepsilon_{kH+1}) = \sigma^2(1 + \rho^2) > \sigma^2$, то есть при $H > 1$ существуют точки с разной дисперсией, если же $H = 1$; то $x_t = \varepsilon_t$ – белый шум.

Замечание: если рассматривать вариант, где H может принимать бесконечные значения, то при $H = \infty$: $x_t = \rho x_{t-1} + \varepsilon_t$ – $AR(1)$ процесс $\Rightarrow$ стационарен.

5. **Условие:**
Рассмотрим следующую VAR(2)-модель.

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \phi_{11}^{(1)} & \phi_{12}^{(1)} \\ \phi_{21}^{(1)} & \phi_{22}^{(1)} \end{pmatrix} \cdot \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \phi_{11}^{(2)} & \phi_{12}^{(2)} \\ \phi_{21}^{(2)} & \phi_{22}^{(2)} \end{pmatrix} \cdot \begin{pmatrix} y_{1,t-1} \\ y_{2,t-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}$$

(a) Кратко опишите, в чём отличия обычных IRF и ортогональных IRF?

(b) В чём отличия интерпретации обычных IRF и кумулятивных?

(c) Одним из преимуществ VAR(2) является тот факт, что через её оценку можно проводить тест Гранжера. Гипотеза: «y_2 является каузальным для y_1». Сформулируйте тест Гранжера относительно коэффициентов $\phi_{ij}^{(p)}$, выпишите нулевую и альтернативную гипотезы.

Решение:

(a) Ортогональные IRF устраняют корреляцию между шоками, что позволяет анализировать реакцию системы на чистый шок одной переменной. Обычные IRF не устраняют корреляцию ошибок, поэтому шок в одной переменной может включать влияние других переменных.

(b) Обычные IRF показывают чистую реакцию на шок через k шагов. Кумулятивные IRF отображают накопленный эффект от шока, то есть суммарный эффект по всем периодам.

(c) Тест Гранжера для модели имеет вид:

$$H_0 : \phi_{12}^{(1)} = \phi_{(12)}^{(2)} = 0 \text{ против } H_1 : \phi_{12}^{(1)} \neq 0 \vee \phi_{12}^{(2)} \neq 0$$

6. **Условие:**
Рассмотрим модель парной регрессии с AR(1) ошибками:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t,\ \varepsilon_t = \rho\varepsilon_{t-1} + u_t,\ u_t \sim \mathcal{N}(0, \sigma_u^2)$$

где $|\rho| < 1$, а x_t – экзогенная переменная.

(a) Найдите распределение ε_1

(b) Найдите распределение $p(y_1|x_1)$

(c) Найдите распределение $p(y_t|y_{t-1}, x_t, x_{t-1})$

(d) Запишите полное правдоподобие $p(y_1, \ldots, y_T|x_1, \ldots, x_T)$

Решение:

(a) Процесс $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$ стационарен, откуда $\mathbb{V}ar(\varepsilon_t) = const$, при этом $\varepsilon_t = \rho\varepsilon_{t-1} + u_t = \rho(\rho\varepsilon_{t-2} + u_{t-1}) + u_t = \rho^k\varepsilon_{t-k} + \rho^k u_{t-k} + \ldots + \rho u_{t-1} + u_t$. Отметим, что u_t – некоррелированные нормальные случайные величины, откуда по свойствам гауссовских векторов, они независимы:

$$\rho^k u_{t-k} + \ldots + \rho u_{t-1} + u_t \sim \mathcal{N}(0; \sigma_u^2 \sum_{i=0}^{k} \rho^{2i})$$

Устремляя $k \to \infty$ и используя $|\rho| < 1$, получаем, что $\sum\limits_{k=0}^{\infty} \rho^k u_{t-k} \sim \mathcal{N}(0, \frac{\sigma_u^2}{1-\rho^2})$. Отметим, что $\mathbb{E}\varepsilon_{t-k} = 0$ и $\mathbb{V}ar(\varepsilon_{t-k}) = const$, откуда $\rho^k\varepsilon_{t-k} \xrightarrow{a.s.} 0$. Таким образом, переходя к пределу, в исходном равенстве для ε_t получаем, что $\varepsilon_t \sim \sum\limits_{k=0}^{\infty} \rho^k u_{t-k} \sim \mathcal{N}(0, \frac{\sigma_u^2}{1-\rho^2})$, откуда и $\varepsilon_1 \sim \mathcal{N}(0, \frac{\sigma_u^2}{1-\rho^2})$.
Заметим, что характеристики распределения можно найти из стационарности процесса. Нормальность распределения ε_t будет следовать из рассуждений, повторяющих предельный переход, и вида u_t, далее:

$$\mathbb{E}\varepsilon_t = \mathbb{E}(\rho\varepsilon_{t-1} + u_t) = \rho\mathbb{E}\varepsilon_{t-1} + 0 \Rightarrow \mu = \rho\mu, |\rho| < 1 \Rightarrow \mu = 0$$

$$\mathbb{V}\varepsilon_t = \mathbb{V}(\rho\varepsilon_{t-1} + u_t) = \rho^2\mathbb{V}\varepsilon_{t-1} + \sigma_u^2 \Rightarrow \sigma^2 = \rho^2\sigma^2 + \sigma_u^2 \Rightarrow \sigma^2 = \frac{\sigma_u^2}{1-\rho^2}$$

(b) $y_1 = \beta_0 + \beta_1 x_1 + \varepsilon_1 \Rightarrow$ нам известно всё, кроме ε_1, а про него знаем распределение, откуда $p(y_1|x_1) \sim \mathcal{N}(\beta_0 + \beta_1 x_1; \frac{\sigma_u^2}{1-\rho^2})$.

(c) $\left.\begin{array}{l} y_t = \beta_0 + \beta_1 x_t + \varepsilon_t = \beta_0 + \beta_1 x_t + \rho\varepsilon_{t-1} + u_t \\ y_{t-1} = \beta_0 + \beta_1 x_{t-1} + \varepsilon_{t-1} \Rightarrow \varepsilon_{t-1} = y_{t-1} - \beta_0 - \beta_1 x_{t-1} \end{array}\right\} \Rightarrow y_t = \beta_0 + \beta_1 x_t + \rho(y_{t-1} - \beta_0 - \beta_1 x_{t-1}) + u_t$
В полученной формуле мы знаем всё, кроме u_t, которое и будет порождать распределение:

$$p(y_t|y_{t-1}, x_{t-1}, x_t) \sim \mathcal{N}(\beta_0(1-\rho) + \beta_1(x_t - \rho x_{t-1}) + \rho y_{t-1}; \sigma_u^2)$$

(d) Используя chain rule, перепишем правдоподобие:

$$p(y_1, \ldots, y_T|x_1, \ldots, x_T) = p(y_1|x_1, \ldots, x_T) \cdot p(y_2|x_1, \ldots, x_T, y_1) \cdot \ldots \cdot p(y_T|x_1, \ldots, x_T, y_1, \ldots, y_{T-1})$$

Для y_1 нет зависимости от предыдущих y_t, поэтому $y_1 = \beta_0 + \beta_1 x_1 + \varepsilon_1$, откуда $y_1 \sim \mathcal{N}(\beta_0 + \beta_1 x_1, \frac{\sigma_u^2}{1-\rho^2})$, то есть:

$$p(y_1|x_1, \ldots, x_T) = p(y_1|x_1) = \frac{\sqrt{1-\rho^2}}{\sqrt{2\pi\sigma_u^2}} exp\Big(-\frac{(y_1 - \beta_0 - \beta_1 x_1)^2(1-\rho^2)}{2\sigma_u^2}\Big)$$

Для $t > 1$ как в пункте (c) существует зависимость между y_t и y_{t-1} и легко заметить, что y_{t-1}, x_{t-1}, x_t оставляют в y_t единственный источник случайности – u_t, откуда можно просто переписать формулу из предыдущего пункта:

$$p(y_t|x_1, \ldots, x_T, y_1, \ldots, y_{t-1}) = p(y_t|x_{t-1}, x_t, y_{t-1}) =$$

$$= \frac{1}{\sqrt{2\pi\sigma_u^2}} exp\Big(-\frac{\big(y_t - \beta_0(1-\rho) - \beta_1(x_t - \rho x_{t-1}) - \rho y_{t-1}\big)^2}{2\sigma_u^2}\Big)$$

Перемножая получившиеся скобки, получаем:

$$p(y_1, \dots, y_T|x_1, \dots, x_T) = p(y_1|x_1, \dots, x_T) \cdot p(y_2|x_1, \dots, x_T, y_1) \cdot \ldots \cdot p(y_T|x_1, \dots, x_T, y_1, \dots, y_{T-1}) =$$

$$= \frac{\sqrt{1-\rho^2}}{\sqrt{2\pi\sigma_u^2}} exp\Big(-\frac{(y_1 - \beta_0 - \beta_1 x_1)^2(1-\rho^2)}{2\sigma_u^2}\Big) \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\sigma_u^2}} exp\Big(-\frac{\big(y_t - \beta_0(1-\rho) - \beta_1(x_t - \rho x_{t-1}) - \rho y_{t-1}\big)^2}{2\sigma_u^2}\Big) =$$

$$= (2\pi)^{-\frac{T}{2}} \cdot \sigma_u^{-T} \cdot exp\left(\frac{(y_1 - \beta_0 - \beta_1 x_1)^2(1-\rho^2) + \sum\limits_{t=2}^{T} \big(y_t - \beta_0(1-\rho) - \beta_1(x_t - \rho x_{t-1}) - \rho y_{t-1}\big)^2}{2\sigma_u^2}\right)$$