1. Body positive Zmey Gorynych worries about his weight but loves eating young beautiful princesses. Each year he makes weight measurements. He is very happy when his weight is between 1000 and 2000 kilograms. Otherwise he is unhappy and does not remember the exact measurement results.

   Here are his recordings for the last 5 years $y = (\leq 1000, 1200, 1500, \geq 2000, 1800)$.

   Assume that his weight measurements are uniform on $[0, a]$ and independent.

   (a) [7] Estimate unknown a using maximum likelihood.

   (b) [3] Estimate the probability that the next year Zmey Gorynych will be happy about his weight.

2. Consider the following equations for the process x_t, y_t:

$$\begin{cases} x_t = 0.5y_{t-1} + 0.25x_{t-2} - 0.5y_{t-2} + u_{1t}, \\ y_t = -0.5x_{t-1} + 0.5y_{t-1} + 0.25x_{t-2} + 0.5y_{t-2} + u_{2t}, \end{cases}$$

   where u_{1t} and u_{2t} are independent white noise processes.

   (a) [2] Represent this system of equations in the $VAR(1)$ form for 4-dimensional process.

   (b) [3] Are there any stationary solutions?

   (c) [3] Derive the VECM (Vector Error Correction Model) representation of these equations.

   (d) [2] Write out the cointegrating vector.

3. Consider the multinomial logit model with 3 alternatives a, b and c. Here we omit index i for observations and just write y instead of y_i for simplicity. The variables D_a, D_b and D_c are «disutilities» of the alternatives. The alternative with the lowest disutility is chosen by individiual.

$$\begin{cases} v_a, v_b, v_c \sim \text{Expo}(\lambda = 1) \\ D_a = v_a/2, \quad D_b = v_b/3, \quad D_c = v_c/4 \\ y = \begin{cases} a, \text{ if } D_a = \min\{D_a, D_b, D_c\} \\ b, \text{ if } D_b = \min\{D_a, D_b, D_c\} \\ c, \text{ if } D_c = \min\{D_a, D_b, D_c\} \end{cases} \end{cases}$$

   (a) [3] What is the distribution of D_a, D_b and D_c?

   (b) [2] What is the distribution of $\min\{D_a, D_b, D_c\}$?

   (c) [3] Derive the probability $\mathbb{P}(y = a)$.

   (d) [2] Derive the probability $\mathbb{P}(y = a \mid y \neq c)$.

4. Consider the following time series model for $\{y_t\}_{t=1}^T$:

$$y_t = \alpha + \beta t + u_t, \quad t = 1, \ldots, T$$

$$\text{with} \quad u_t = \rho u_{t-1} + \varepsilon_t \quad \text{and} \quad |\rho| \leq 1$$

where ε_t is an i.i.d. error with distribution $(0, \sigma_\varepsilon^2)$ that is uncorrelated with anything in the past (white noise).

   (a) [4] Show that y_t is trend stationary when $|\rho| < 1$ [ie there is a trend stationary solution].

   (b) [3] Show that y_t is difference stationary when $\rho = 1$.

   (c) [3] Discuss the importance of distinguishing between trend stationary and difference stationary processes.

5. Consider the equation

$$y_t = y_{t-1} + u_t - 1.5u_{t-1} + 0.5u_{t-2},$$

where (u_t) is a white noise process.

   (a) [4] Write this equation in the form $A(L)y_t = B(L)u_t$ with factored polynomial $B(L)$.

   (b) [2] Provide an example of a stationary solution of the equation if possible.

   (c) [2] Provide an example of a difference stationary solution of the equation if possible.

   (d) [2] Provide an example of a non-stationary solution that is not difference stationary if possible.

6. Let us consider an analysis on recidivism (probability of re-arrest) among a group of young men in California who have at least one arrest prior to 1986. The dependent variable, `arr86`, is equal to unity if the man was arrested at least once during 1986, and zero otherwise. The standard errors are reported in parentheses.

	OLS A	OLS B	Logit A	Logit B	Logit B Marginal Effect
pcnv	$-.152$ (.021)	$-.162$ (.021)	$-.880$ (.122)	$-.901$ (.120)	$-.176$ (.023)
avgsen	.005 (.006)	.006 (.006)	.027 (.035)	.031 (.034)	.006 (.007)
tottime	$-.003$ (.005)	$-.002$ (.005)	$-.014$ (.028)	$-.010$ (.027)	$-.002$ (.005)
ptime86	$-.023$ (.005)	$-.022$ (.005)	$-.140$ (.031)	$-.127$ (.031)	$-.025$ (.006)
qemp86	$-.038$ (.005)	$-.043$ (.005)	$-.199$ (.028)	$-.216$ (.028)	$-.042$ (.005)
black	.170 (.024)	–	.823 (.117)	–	–
hispan	.096 (.021)	–	.522 (.109)	–	–
constant	.380 (.019)	.380 (.019)	$-.464$ (.095)	$-.169$ (.084)	
R^2	.068	.047			
$\log L$			-1512.35	-1541.24	

`pcnv` is the proportion of prior arrests that led to a conviction,
`avgsen` is the average sentence served from prior convictions,
`tottime` is the months spent in prison since age 18 prior to 1986,
`ptime86` is months spent in prison in 1986,
`qemp86` is the number of quarters the man was legally employed in 1986,
`black` and `hispan` are two race dummies (white the excluded dummy).

(a) [2] Using the OLS B results, what is the estimated effect on the probability of arrest if `pcnv` goes from 0.25 to 0.75 holding everything else constant?

(b) [2] It is argued that the linear probability model is not appropriate for explaining the binary variable `arr86` and a logit regression model has been estimated. Explain how the Logit estimates are obtained.

(c) [3] Using the Logit model results, discuss whether `black` and `hispan` are jointly significant. Clearly indicate the null and alternative hypothesis, the test statistic and the rejection rule.

(d) [3] Using the Logit B results, how would you obtain the estimated effect on the probability of arrest if `pcnv` goes from 0.25 to 0.75 for a white man, with characteristics `avgsen=1`, `tottime=1`, `ptime86=0` and `qemp86=2`. A clear explanation of what calculations are required is sufficient.