Problem 1

For time series given in the table:

t	1	2	3	4	5	6	7	8	9
y_t	85	81	78	72	69	70	64	61	56

The trend model is built as $\widehat{y_t} = 87.8 - 3.4t$. Estimate the model adequacy and accuracy.

Solution

t	y_t	$\widehat{y_t}$	$\varepsilon_t = y_t - \widehat{y_t}$	peak point	ε_t^2	$\varepsilon_t - \varepsilon_{t-1}$	$(\varepsilon_t - \varepsilon_{t-1})^2$	$\lvert\varepsilon_t\rvert : y_t \cdot 100$
1	85	84	0	-	0	- 0	0	0
2	81	81	0	*1*	0	0	0	0
3	78	77	0	1	0	- 2	6	0
4	72	74	- 2	1	4	0	0	2
5	69	70	- 1	0	2	4	19	2
6	70	67	2	1	7	- 2	6	3
7	64	63	0	1	0	0	0	0
8	61	60	0	1	0	- 1	2	0
9	56	57	-1	-	1	- 0	0	1
Total	636	636.3	-0.3	6	15.51		35.26	13.29

First of all, let's form a residual sequence, that consist of differences between real and predicted values (it is given in 4-th column). Next let's check if the levels of residual sequence are stochastic. Peak points are marked in column 5, there are 6 peak points. Next we must check if the inequality

$$p > \left[\bar{p} - 1.96\sqrt{\sigma_p^2}\right]$$

holds, where $\bar{p} = \frac{2}{3}(n-2), \sigma_p^2 = \frac{16n-29}{90}$. Substituting n we see that right part of inequality is equal to 2, right (number of peak point) equals 6, so inequality holds, so residual series is stochastic. Now let's check if the residual sequence is normally distributed. Let's use RS-criteria to check this. $R = \varepsilon_{\max} - \varepsilon_{\min} = 2.7 - (-2.1) = 4.8$, standard deviation $S_{\hat{y}} = \sqrt{\frac{\varepsilon_t^2}{n-1}} = \sqrt{\frac{15.51}{8}} = 1.39$. Next, $RS = \frac{4.8}{1.39} = 3.45$. This value is between lower and upper levels of table values of this criteria (for n=10 and $\alpha = 0.05$ these levels are 2.7 and 3.7). So, residuals are distributed normally.

Now let's check equality to 0 of expectation. It is equal to $-\frac{0.3}{9} \approx -0.03$, so we can confirm this without Student's statistics.

Now let's check independency of residual levels (autocorrelation absence). Let's use Durbin-Watson statistics. Using columns in table value of the criterion is equal to $d = \frac{35.26}{15.51} = 2.27 > 2$. This means that autocorrelation is negative, so we must transform the criterion. $d' = 4 - d = 4 - 2.27 = 1.73$. Comparing this with 2 table values, that for linear model are equal to $d_1 = 1.08, d_2 = 1.36$. Since $d_2 < d < 2$ we conclude residual sequence levels are independent.

So, residual sequence satisfy all properties of stochastic component of time series, so linear model is adequate.

Now let's characterize model accuracy. Let's use mean approximation error. $\bar{\varepsilon}_{relative} = \frac{13.29}{9} = 1.48\%$<5%. So, model is sufficiently accurate.