Homework 3 (due: Fr, Oct. 13)

The purpose of Problems 1 and 2 is to investigate the dependence of the time value $V_n$ of a call option at the money with a given strike price as a function of the time $n$ to expiration. You will analyze market data using (simple) linear regression (data fitting or method of least squares). This will be briefly reviewed in the TA session, but I assume that you know this.

(1) The following table gives the prices on 9/28/16 for call options on Wells Fargo (ticker: WFC) at strike price $45$ for various expiration dates. The last price price for WFC was $45.31$ and so the options are at the money. The table shows the days $n$ until expiration and the price of the option. The prices shown are averages of the last bid/ask prices. As the spreads were generally small in the range of 10–30 cents, this provides a good approximation of the market value of the options.

<table>
<thead>
<tr>
<th>days $n$ until exp.</th>
<th>price of option</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>1.02</td>
</tr>
<tr>
<td>22</td>
<td>1.15</td>
</tr>
<tr>
<td>29</td>
<td>1.25</td>
</tr>
<tr>
<td>36</td>
<td>1.31</td>
</tr>
<tr>
<td>50</td>
<td>1.43</td>
</tr>
<tr>
<td>78</td>
<td>1.72</td>
</tr>
<tr>
<td>113</td>
<td>2.06</td>
</tr>
<tr>
<td>204</td>
<td>2.63</td>
</tr>
<tr>
<td>260</td>
<td>2.89</td>
</tr>
<tr>
<td>477</td>
<td>4.01</td>
</tr>
<tr>
<td>841</td>
<td>5.25</td>
</tr>
</tbody>
</table>

(a) Compile a table with the time values of the options.

(b) A reasonable assumption seems that the time value $V_n$ for an option in our table with $n$ days until expiration follows an approximate law of the form

$$V_n = An,$$

where $A$ is some coefficient. To justify this linear relation, one could argue that if an option $O_1$ has twice as long until expiration as another option $O_2$ with the same strike, then the time value of $O_1$ should be twice the time value of $O_2$.

Analyze the data by using linear regression (method of least squares) to find the coefficient $A$ that gives the best data fit for our assumed law $V_n = An$.

Hint: You can do the computations by hand (tedious!) or use a spreadsheet such as Excel or LibreOffice (free to download from the internet) that has built-in
procedures for linear regression of data (recommended!). In the latter case, add a printout with your data.

(c) Using the coefficient computed in (b), compile a table of the values $A_n$ that are supposed to approximate the values $V_n$ and compare them with our real data. What do you think: is our model $V_n = A_n$ viable?

(2) (a) Maybe a more realistic model in Problem 1 is a relation of the form $V_n = B n^\kappa$, where $B > 0$ and $\kappa > 0$ are some coefficients. Is $B$ or $\kappa$ more important here for the general characteristics of our model?

(b) If we take (natural) logarithms in (a), then we obtain

$$\log V_n = b + \kappa \log n,$$

where $b = \log B$. In other words, in this model there should be a linear relation between $\log V_n$ and $\log n$.

Use linear regression to find the constants $b$ and $\kappa$ here that give the best fit for the data.

(c) Based on the numerical evidence in (b), what seems to be a plausible and simple assumption for the “true” value of $\kappa$?

(d) Use the best fit value $B = e^b$ obtained in (e) and the “true” value of $\kappa$ from (f), to compile a list of the values $B n^\kappa$. Compare them with the real data for $V_n$. What is your assessment?

(e) Summarize your numerical findings in a simple rule of thumb: if an option $O_1$ at the money with the same strike as another option $O_2$ has a time value that is twice as big as the time value of $O_1$, then the time to expiration of $O_1$ should be $X$ times the time to expiration of $O_2$. What is $X$ here?

(3) (a) Let $X$ be a standard normal random variable. Then we denote by $N$ its cumulative distribution function given by

$$N(x) = \mathbb{P}(X \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$$

for $x \in \mathbb{R}$. The function $N$ cannot be expressed in any simple way by other known functions (such as the exponential function, trigonometric functions, etc.), but it is a built-in function in standard statistical programs and so its values are easy to obtain.

Plot the graph of $N$ in the range $x \in [-3, 3]$.

(b) Give a probabilistic explanation why $N(x)$ is very close to 1 if $x$ is large, and $N(x)$ is close to 0 if $x$ is small (i.e., a large negative number). What is the precise mathematical formulation of these facts in terms of limits?

(c) Show that $N(x) + N(-x) = 1$ for each $x \in \mathbb{R}$. 
(d) Compute $N(0)$ and $N'(0)$.

(e) Let $X$ be a standard normal random variable and $c$ be a real number. Show that then

$$E((X - c)_+) = g(c) := \frac{1}{\sqrt{2\pi}}e^{-c^2/2} - cN(-c).$$

(4) At a certain date a stock has a price $S_0$ (per share). Let us assume that the stock price has neither an upward nor a downward trend and that it goes “sideways”. Then a simplistic model for how the stock price develops over time is that it goes up or down in a random way and that the change each day is by $\pm \sigma$, where $\sigma$ is a (small) fixed amount and a gain or loss of $\sigma$ in the stock price happens with equal probability. We also assume that these day-to-day changes are independent and ignore the effect of interest rates.

(a) We denote the stock price at day $n$ by $S_n$. Based on our model, express $S_n$ in terms of $S_0$ and suitable i.i.d. random variables $Z_1, Z_2, \ldots$.

(b) Based on the central limit theorem our model suggests that $S_n$ for large $n$ has a similar distribution as $c_n + d_nN$, where $N$ is a standard normal random variable, and $c_n$ and $d_n$ are suitable constants. What are $c_n$ and $d_n$ here?

(c) Even though $c_n$ in (b) is non-negative, the random variable $c_n + d_nN$ can take negative values, which of course cannot happen for the stock price $S_n$. We can ignore this problem of our model if $c_n + d_nN$ is rather sharply centered around its mean value $c_n > 0$. Formulate a corresponding condition in terms of $S_n$, $n$, and $\sigma$ that is relevant for the validity of our model!