Lab Report #2: Sums, Mixtures, & Certainty Equivalents
Revised: September 18, 2014

Due at the start of class. You may speak to others, but whatever you hand in should be your own work. Use Matlab where possible and attach your code to your answer.

> **Solution:** Brief answers follow, but see also the attached Matlab program; download the pdf, open, click on pushpin:

1. *Properties of US equity returns.* Gene Fama is one of the giants of modern finance. His coauthor over the past twenty years or so has been Ken French. In this problem, we'll compute sample properties of US equity returns using one of the so-called Fama-French datasets, which French kindly posts on his website. Go to French's website and download the files labelled "Fama/French factors" at this link. Then follow these instructions:

   - Read the enclosed txt file into a spreadsheet program.
   - Strip off the header (the three lines at the top).
   - Label the columns `date xs smb hml rf`.
   - Delete the annual data at the bottom.
   - Save in a convenient format (xls?).

   When you're done, read the whole thing into Matlab using `xlsread` or other command appropriate to the format of your file.

   You now have monthly data going back to 1926 on various equity portfolios, each in its own column. Returns are expressed as percentages. They haven't been annualized: a monthly return of 1.0 in the data corresponds to roughly 12% annually. We'll focus on `xs`, the excess return on a broad-based equity portfolio. Excess return means, in this context, the return minus that on the riskfree bond, `rf`. Also there but not needed for now: `smb` is the return on small firms minus big firms and `hml` is the return on firms with high book-to-market minus those with low book-to-market.

   (a) Use the command `hist` to plot a histogram for `xs`.
   (b) Compute the mean, standard deviation, skewness, and excess kurtosis for the excess return series.
   (c) Does the mean look high or low to you?
   (d) In what respects do these excess returns look normal? In what respects not?

Solution: I did this in Python, partly because it's something new, but mostly because it allows me to load the data in two lines:

```
import pandas.io.data as web
ff = web.DataReader('F-F_Research_Data_Factors', 'famafrench')[0]
```

More at this link, Including histograms, box plots. and summary statistics. See Example 2.

(a) See the link above. I especially like the boxplots. The box gives us the median and the (25,75)th percentiles. The points outside the box are what you might call outliers: extremes in one direction or the other.

(b) The numbers are

Series	Mean	Std Dev	Skewness	Ex Kurtosis
Market	0.649	5.40	0.16	4.48
SMB	0.224	3.23	2.05	17.57
HML	0.394	3.51	1.93	12.87

Note that you have to subtract 3 to get excess kurtosis from the kurtosis command in Matlab.

(c) Hard to say what's high or low, it depends on your perspective. But if we multiply by 12 to get a rough annual number, the excess return on the market is about 7.7%, which seems large to me. The other portfolios have smaller means, but they also have smaller standard deviations.

(d) It's clear that these returns are a long way from normal. There's some evidence of positive skewness (hmmm), but more of excess kurtosis, especially in the SMB and HML portfolios. Both would be close to zero if returns were normal. The histograms look normal-ish, but that's a sign of how hard it is to tell this way.

2. *Sums and mixtures.* More fun with cumulant generating functions (cgfs). The ingredients include two independent normal random variables, $x_1 \sim \mathcal{N}(0,1)$ and $x_2 \sim \mathcal{N}(\theta, 1)$. Do the calculations in Matlab and submit your code with your answer.

(a) Consider the sum $y = x_1 + x_2$. What is its cgf? Its first four cumulants? What are its measures of skewness and excess kurtosis, γ_1 and γ_2?

(b) Now consider the mixture z:

$$z = \begin{cases} x_1 & \text{with probability } 1 - \omega \\ x_2 & \text{with probability } \omega. \end{cases}$$

What is its cgf? Its first four cumulants? What are its measures of skewness and excess kurtosis, γ_1 and γ_2? What determines the sign of γ_1?

(c) In what ways does z differ from a normal random variable?

Solution:

(a) The cgf of the sum is the sum of the cgf's:

$$k_y(s) = k_1(s) + k_2(s) = s(0 + \theta) + s^2(1 + 1)/2.$$

The form of the cgf should be familiar, we derived it in class. Since it's quadratic, y must be normal.

The cumulants follow from differentiating (or just looking, really): $\kappa_1 = \theta$ (the mean), $\kappa_2 = 2$ (the variance), $\kappa_3 = 0$, and $\kappa_4 = 0$. The last two are zero, because the cgf is quadratic in s, so all the derivatives after the second one are zero. Therefore $\gamma_1 = \gamma_2 = 0$ as well.

(b) Mixtures are a little different. The mgf is the weighted average of the mgf's of the components:

$$h_y(s) = (1 - \omega)h_1(s) + \omega h_2(s) = (1 - \omega)e^{s^2/2} + \omega e^{\theta s + s^2/2}.$$

The cgf is the log of this. The first four cumulants are $\kappa_1 = \omega\theta$ (the mean), $\kappa_2 = 1 + \omega(1 - \omega)\theta^2$ (the variance), $\kappa_3 = \omega\theta^3(1 - 3\omega + 2\omega^2)$, and $\kappa_4 = 3\omega(1 - \omega)\theta^4[1 - 6\omega(1 - \omega)]$.

The striking thing here is that we're getting nonzero third and fourth cumulants from normal components. The details are a mess so I'll skip them. The sign of γ_1 (skewness) is the same as θ if ω is small.

(c) Skewness and excess kurtosis are a sign that z is not normal.

---

3. *Concave and convex functions.* Consider the following functions defined over positive values of x:

   (a) $f(x) = \log x$
   (b) $f(x) = \exp(x)$
   (c) $f(x) = x^\alpha$ for $\alpha = 2$
   (d) $f(x) = -x^\alpha$ for $\alpha = 2$
   (e) $f(x) = x^\alpha$ for $\alpha = 1$

   For each one:

   - State whether it's concave, convex, or something else.
   - State whether $E[f(x)]$ is greater or less than $f[E(x)]$.
   - Verify your previous answer by computing $E[f(x)]$ and $f[E(x)]$ for the Bernoulli random variable

   $$x = \begin{cases} 100 & \text{with probability } 1/2 \\ 200 & \text{with probability } 1/2. \end{cases}$$

Solution: This is a check on Jensen's inequality.

- (a) Concave, (b) convex, (c) convex, (d) concave, (e) linear (on the edge between concave and convex).

- If concave $E[f(x)] < f[E(x)]$. If convex the reverse. If linear, they're equal.

- Summary:

	(a)	(b)	(c)	(d)	(e)
$E[f(x)]$	4.95	10^{86}	25,000	$-25{,}000$	150
$f[E(x)]$	5.01	10^{65}	22,500	$-22{,}500$	150

4. *Certainty equivalents for quadratic utility.* There's a long history of risk preference with quadratic utility:

$$u(c) \;\; = \;\; c - c^2.$$

It's seldom used these days, but is a convenient starting point for thinking about risk. Let us say, to be concrete, that c is random with mean μ and variance σ^2.

(a) What is expected utility? How does it depend on the distribution of consumption outcomes?

(b) What is the certainty equivalent?

(c) How does the certainty equivalent compare to expected consumption?

Solution: I'm going to change the notation so that consumption has a mean of κ_1 and a variance of κ_2. That way I can use μ for the certainty equivalent.

(a) If we take the expectation, we get

$$E[u(c)] \;\; = \;\; E(c) - E(c^2) \;\; = \;\; \kappa_1 - (\kappa_1^2 + \kappa_2).$$

The point is that only the mean and variance matter. That was thought to be a strength when people first started doing things like this, but is now thought to be a severe limitation. It's far more common now to see power utility.

It wasn't mentioned, but this is only sensible for a range of c's: we want u increasing, which requires $1 - 2c > 0$ or $c < 1/2$. This makes the rest of the problem tricky.

(b) The certainty equivalent μ is the solution to

$$u(\mu) \;\; = \;\; \mu - \mu^2 \;\; = \;\; \kappa_1 - (\kappa_1^2 + \kappa_2).$$

This is quadratic, so it has two solutions. Skip this.

(c) Skip this.