Financial Econometrics

Nicky Grant

EC5609: Problem Set 1

Semester 1 2021/22

1. For the standard uniform random variable $X \sim Unif[0,1]$ where the probability density function is $f(x) = 1$ for $0 \leq x \leq 1$ for 0 otherwise calculate the mean, variable, skewness and kurtosis.

   Mean: $E[X] = \int_0^1 1.xdx = [x^2/2]_0^1 = 1/2$
   Variance: $\sigma^2 := E[X - E[X]]^2 = \int_0^1 (x - 1/2)^2 = [\frac{1}{3}(x - 1/2)^3]_0^1 = 1/3(1/2^3 - -1/2^3) = 2/3(1/8) = 1/12$
   Skewness: $E[(X - E[X]/\sigma)^3] = 0$ since the distribution is symmetrical around mean 1/2
   Kurtosis: $E[((X - E[X])/\sigma)^4] = (\sqrt{12})^4(\int_0^1 (x - 1/2)^4) = (12)^2[1/5(x - 1/2)^5]_0^1 = (12)^2 * 1/5(1/2^2 + 1/2^5) = 144 * 2/5 * 1/64 = 144/160 = 9/10$

2. For two random variables X and Y, provide a proof of the Law of Iterated Expectations, namely that

   $$E_Y[Y] = E_X\left[E_{Y|X}[Y|X]\right] \quad (1)$$

   you may assume that X,Y are continuously distributed (likewise the proof can be shown for discrete variables).

   $E_X\left[E_{Y|X}[Y|X]\right] = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} yf(y|x)dy\right] f(x)dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(y|x)f(x)dydx$
   Note that by Bayes Rule $f(y|x)f(x) = f(y,x)$
   $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(y|x)f(x)dydx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(y,x)dydx = \int_{-\infty}^{\infty} y(\int_{-\infty}^{\infty} f(y,x)dx)dy$
   Where by definition we get the marginal density of y by integrating the joint density function over all values x can take, i.e.$f(y) = (\int_{-\infty}^{\infty} f(y,x)dx$
   $\int_{-\infty}^{\infty} y(\int_{-\infty}^{\infty} f(y,x)dx)dy = \int_{-\infty}^{\infty} yf(y)dy = E_Y[Y]$

3. The weekly returns of the Apple stock for 1/1/01 to 12/2/01 are (0.101,0.050, 0.135, 0.003,0.0543,-0.072, -0.007, -0.007,-0.009).
   Using this data set of size $T = 9$ answer the following (all the below can be done on a calculator, and I would advise you use a calculator in this example).

   (a) Calculate the sample mean and variance.
   Sample Mean $\bar{x} = \frac{1}{T}\sum_{t=1}^{T} x_t$ in our case $T = 9$ and we substitute in the 9 samples values of returns $\bar{x} = 0.02849$
   Sample Variance: $\frac{1}{T}\sum_{t=1}^{T}(x_t - \bar{x})^2 = \frac{1}{9}\sum_{t=1}^{T}(x_t - 0.02849)^2 = 0.398$

   (b) Calculate the sample autocorrelation at lag 1 and 2.
   Sample correlation at lag k $\hat{\rho}(k) = \frac{1}{T-k}\sum_{t=k+1}^{T}(x_{t-k} - \bar{x})(x_t - \bar{x})$ which we evaluate at $T = 9$ for $k = 1$ and $k = 2$ noting we lose the first (the first two) observations for correlations at lag 1 and 2 respectively. Plug in $\bar{x}$ from above and the remaining elements from the data vector and sum (you need to be comfortable calculating these by hand for the class test and exam).
   $\hat{\rho}(1) = 0.1178$ and $\hat{\rho}(2) = 0.586$

(c) Perform a 95% two sided hypothesis test that the population correlation at lag 1 is zero (state all the steps, the test statistic and distribution and any assumptions made).
Test stat: $\sqrt{9} * 0.1178 = 0.354$
Under H_0 (under the stronger hypothesis of i.i.d data) then the test stat is drawn from a $N(0,1)$ approximately (when T large). The 95% critical values for a two sided test are $+/-1.96$ and hence our test stat is within this range so there is insufficient evidence to reject the null hypothesis that the true population correlation at lag 1 is zero.

(d) Perform a 95% two sided joint hypothesis test that the population correlation at lag 1 and 2 are both zero (state all the steps, the test statistic and distribution and any assumptions made).
Q-Stat $\sum_{t=1}^{2}(\sqrt(T)\hat{\rho}(k))^2 = 3.8183$. Under the null the the stat is approximately χ_2^2 distributed. The p-value for this test is 0.1482, and so we cannot find evidence to reject the null at the 5% level. (Alternatively you may find the 95% critival value from the statistical tables, in this case 5.991)

(e) Why may we be sceptical of inference from the results in (c),(d).

4. Let $\mathcal{F}_t$ denote the information set of all past prices and $\mathcal{G}_t$ the information set with all past prices and all public information.

(a) Provide a statement of the weak and semi strong form of the efficient market hypothesis (EMH) and explain this result.
See the definition in Lecture 2.

(b) Show that the semi strong form of EMH implies the weak form.
Follows by the law of iterated expectations. $E[R_T|G_t] = R_T^*$ for all $T > t$ and all public information sets G_t. Since the G_T contains F_T then the average of R_T for an possible past prices is on average R_T^* in every possible world of public information- hence the average R_T conditional on F_t is R_T^* as no matter what the public info is the average of returns is the normal returns for all possible prices.

(c) Show that the weak form EMH implies returns in different time periods are uncorrelated when the normal rate of return is constant.
$R_T^* = R$ for all T (constant normal returns) hence WF EMH implies $E[R_T|F_t] = R$ for all $T > t$ hence $E[R_T - R|F_t] = 0$. Then take any $R_{T-j} - R$ for $j > 0$ then we can multiply both sides of $E[R_T - R|F_t] = 0$ by $(R_{T-j} - R)$ and note that $T - j < T$ and hence R_{T-j} is known and fixed and so $E[(R_T - R)(R_{T-j} - R|F_t] = 0$. Then by law of iterated expectations $E[(R_T - R)(R_{T-j} - R)] = 0$ noting also $E[R_T] = R$ for all T by law of iterated expectations (namely take expectations on both sides of $E[R_T|F_t] = R$). Hence $Cov(R_T, R_{T-j}) = E[(R_T - R)(R_{T-j} - R)] = 0$ for all $T, j > 0$ so that if the normal rate of return is constant EMH implies returns are uncorrelated in different time periods.

(d) If normal rate of return is not constant, does the result in (c) hold, and if not explain why.
If the normal rate of return isn't constant we have $E[R_T - R_T^*|F_t] = 0$ and this implies $Cov(R_T - R_T^*, R_{T-j} - R_{T-j}^*) = 0$. In general $Cov(R_T - R_T^*, R_{T-j} - R_{T-j}^*)$ depends on the covariance of R_T^* and so $Cov(R_T, R_{T-j})$ does not equal $Cov(R_T - R_T^*, R_{T-j} - R_{T-j}^*)$.

5. Below is the joint distribution of returns R_t for $t = 1$ and $t = 2$ that take 3 values, 0,1 or 2 percent.

(a) Derive the marginal distributions of R_1 and R_2.

R1: $Pr(R_1 = 0) = 0.18$ $Pr(R_1 = 1) = 0.41$ $Pr(R_1 = 2) = 0.31$

R2:$Pr(R_2 = 0) = 0.34$ $Pr(R_2 = 1) = 0.36$ $Pr(R_2 = 2) = 0.3$

(b) Derive the mean and variance of R_1, R_2.

Mean of R1: $E[R_1] = 0 * 0.18 + 1 * 0.41 + 2 * 0.31 = 1.03$

Mean of R2: $E[R_2] = 0 * 0.34 + 1 * 0.36 + 0.3 * 2 = 0.96$

(c) Find $Cov(R_1, R_2)$

$Cov(R_1, R_2) = E[R_1 R_2] - E[R_1]E[R_2] = 1.11 - 0.96 * 1.13 = 0.0252$ where $E[R_1 R_2] = 1 * 1 * 0.11 + 1 * 2 * 0.15 + 2 * 1 * 0.19 + 2 * 2 * 0.08 = 1.11$

(d) Find the conditional distribution of R_1 given $R_2 = 0, 1$ or 2 respectively.

$R_1|R_2 = 0$: takes probability 0,1,2 with conditional prob $0.05/0.34 = 0.147$ $0.21/0.34 = 0.618$ and $0.08/0.34 = 0.235$ respectively (all using Bayes Rule)

$R_1|R_2 = 1$ takes probability 0,1,2 with conditional prob $0.1/0.36 = 0.277$, $0.11/0.36 = 0.306$ and $0.15/0.36 = 0.417$ respectively

$R_1|R_2 = 2$ takes probability 0,1,2 with conditional prob $0.03/0.3 = 0.009$ $0.19/0.3 = 0.63333$ and $0.08/0.3 = 0.26667$ respectively

(e) What does (d) tell us about the validity of the efficient market hypothesis?

We can see that R_1 and R_2 are dependent on each other as the outcomes of each impact the probabilities of

		R_2		
		0	1	2
R_1	0	0.05	0.1	0.03
	1	0.21	0.11	0.19
	2	0.08	0.15	0.08