ECON-UB 233

Dave Backus @ NYU

Lab Report #1: Moments & Cumulants

Revised: September 11, 2013

Due at the start of class. You may speak to others, but whatever you hand in should be your own work. Please include your Matlab code.

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

1. *Moments of the standard normal.* This should be review, but will get you started with moments and moment generating functions.
   (a) What is the probability density function of the standard normal?
   (b) What is its moment generating function (mgf)? (Don't derive it, just write it down.)
   (c) Use Matlab to differentiate the mgf and find the first two moments. Are they raw or central moments (μ'_j or μ_j)?
   (d) Find the third and fourth moments the same way. What are they?

Solution:

(a) The word "standard" here means $\mu = 0$ and $\sigma = 1$, so we have $p(x) = (2\pi)^{-1/2} \exp(-x^2/2)$.

(b) $h(s) = \exp(s^2/2)$.

(c) The first two moments are zero and one. They're both raw and central moments, because the mean is zero.

(d) The third and fourth moments are zero and three.

2. *Sample moments of the standard normal.* It's often helpful to experiment with artificial test problems, just to remind ourselves how the code works. Here we compute sample moments of artificial data generated in Matlab and verify that calculations of various moments do what we think they do.

   This generates the data we'll use:

```
format compact      % single-spacing of output
nobs = 1000;        % number of observations
rng('default');     % sets "seed" so you can replicate the output
x = randn(nobs,1);  % generates a vector of standard normals
```

These commands generate "pseudo-random" numbers from a standard normal distribution and put them in the vector x. (Standard normal means normal with mean equal to zero and variance equal to one.) As always, you can find out what Matlab commands do by typing `help command` at the prompt; for example, `help rng` or `help randn`.

(a) Our first check is to see if the sample moments correspond, at least approximately, to our knowledge of normal random variables. For example, use the commands:

```
xbar = mean(x)
moments = mean([(x-xbar).^2 (x-xbar).^3 (x-xbar).^4])
```

What do you get? How do your calculations compare to the analogous moments of the standard normal distribution?

(b) Our second check is on the Matlab commands `mean(x)`, `std(x)`, `skewness(x)`, and `kurtosis(x)`. How do they compare to the sample moments you just computed? Are they exactly the same, almost the same, or completely different? Do you know why?

Solution:

(a) The moments are close to what the distribution implies: mean one, variance (and standard deviation) one, skewness zero, and kurtosis three. The small differences reflect sampling variability. You can see this by increasing the sample size, which typically makes the sample moments closer to the "population" moments. The one difference is in the variance and standard deviation: the Matlab functions divide the sum of squared deviations by the number of observations **minus one**, rather than just the number of observations. It's a small difference, to be sure.

(b) The `skewness` and `kurtosis` commands give exactly the same answers as our own calculations, which assures us that they do what we want them to do.

3. *Cumulants of Bernoulli random variables.* Consider a random variable x that equals δ (an arbitrary number) with probability ω (a number between zero and one) and 0 with probability $1-\omega$. We'll use its cumulant generating function (cgf) to find its first four cumulants, representing, respectively, its mean, variance, skewness, and kurtosis. Do the calculations in Matlab and submit your code with your answer.

(a) Verify that this is a legitimate probability distribution.

(b) What is the mean? The variance? The standard deviation?

(c) Derive the moment generating function. (If you're confused about this, apply the definition.) What is the cumulant generating function?

(d) Differentiate the cgf to find the first four cumulants, labelled κ_1 through κ_4. What are the mean and variance?

(e) Derive the standard measures of skewness and excess kurtosis:

$$\begin{aligned}\gamma_1 &= \kappa_3/(\kappa_2)^{3/2} \quad \text{(skewness)} \\ \gamma_2 &= \kappa_4/(\kappa_2)^2 \quad \text{(excess kurtosis)}\end{aligned}$$

How do they depend on ω? δ? What is excess kurtosis when $\omega = 1/2$?

Solution: This the Bernoulli multiplied by δ, so it illustrates the impact of scaling.

(a) Since ω and $1-\omega$ are nonnegative and sum to one, we're ok.

(b) Mean: $\delta\omega$. Variance: $\delta^2\omega(1-\omega)$. Standard deviation: square root of variance.

(c) The mgf is $h(s) = 1-\omega+\omega e^{s\delta}$. The cgf is $k(s) = \log h(s)$. Everything so far should look familiar from class.

(d) The cumulants are

$$\begin{aligned}
\kappa_1 &= \delta\omega \\
\kappa_2 &= \delta^2\omega(1-\omega) \\
\kappa_3 &= \delta^3\omega(1-\omega)(1-2\omega) \\
\kappa_4 &= \delta^4\omega(1-\omega)[1-6\omega(1-\omega)].
\end{aligned}$$

Note the scaling: κ_j includes δ^j. Other than scaling, we've seen the first two before. The third one tells us that skewness depends on the sign of δ and whether ω is greater or less than one half (graph the probabilities against x if this isn't clear). The fourth one depends on $\omega(1-\omega)$. At $\omega = 1/2$, this terms reaches its max of $1/4$, so the overall term is negative, which generates negative excess kurtosis. As ω moves toward zero or one, this term shrinks and excess kurtosis rises.

(e) You'll note that Matlab doesn't do the obvious cancellation of δ's. Once you do, you have

$$\begin{aligned}
\gamma_1 &= \operatorname{sgn}(\delta)(1-2\omega)/[\omega(1-\omega)]^{1/2} \\
\gamma_2 &= 1/[\omega(1-\omega)] - 6.
\end{aligned}$$

There's a subtle issue with γ_1: the magnitude of δ doesn't matter, but its sign ("sgn") does. That shows up in the ratio of $(\delta^2)^{3/2}$ to δ^3 (think about this a minute). What about excess kurtosis? At $\omega = 1/2$, $\gamma_2 = -2$, so there's negative excess kurtosis. Loosely speaking, it has thinner tails than the normal. As ω approaches zero or one, we reverse that. As we increase/decrease ω, we get a distribution with increasing skewness and excess kurtosis.

4. *Exponential and gamma random variables.* Two common distributions for positive random variables are the exponential and the gamma. We say x is exponential and y is gamma if their pdf's are

$$\begin{aligned}
p(x) &= \lambda e^{-\lambda x} \\
p(y) &= y^{\alpha-1}e^{-\beta y}[\beta^\alpha/\Gamma(\alpha)]
\end{aligned}$$

for $x, y \geq 0$. For $p(x)$, the parameter $\lambda > 0$. For $p(y)$, $\alpha, \beta > 0$. The term in brackets at the end is a constant, chosen so that the pdf integrates to one. Γ is the gamma function, which has the property $\Gamma(n) = (n-1)!$.

We're going to demonstrate some connections using their cgf's, which are

$$
\begin{aligned}
k_x(s) &= -\log(1 - s/\lambda) \\
k_y(s) &= -\alpha \log(1 - s/\beta).
\end{aligned}
$$

As usual, do the calculations in Matlab and submit your code with your answer.

(a) What is the skewness γ_1 of a gamma random variable? How does it compare to the skewness of an exponential?

(b) For what choice of parameter values does a gamma random variable have an exponential distribution?

(c) What is the cgf of the sum of n independent exponential random variables? How does it compare to the cgf of a gamma random variable?

(d) What is the cgf of the sum of two independent and identical gamma random variables?

Solution:

(a) Note that the exponential is a special case of the gamma: set $\alpha = 1$ and $\beta = \lambda$. For the gamma, we have

$$
\begin{aligned}
\kappa_1 &= \alpha/\beta \\
\kappa_2 &= \alpha/\beta^2 \\
\kappa_3 &= 2\alpha/\beta^3 \\
\kappa_4 &= 6\alpha/\beta^4,
\end{aligned}
$$

which gives us

$$
\begin{aligned}
\gamma_1 &= 2/\alpha^{1/2} \\
\gamma_2 &= 6/\alpha.
\end{aligned}
$$

When $\alpha = 1$ (the exponential case), we have $\gamma_1 = 2$ and $\gamma_2 = 6$.

(b) Done.

(c) If we sum n independent exponentials, the cgf is the sum, which is the original cgf times n:

$$
k(s) = -n \log(1 - s/\lambda).
$$

This has the same form as the gamma, with $\alpha = n$ and $\beta = \lambda$.

(d) Similarly, if we sum two independent gammas, the cgf is two times the original:

$$
k(s) = -2\alpha \log(1 - s/\beta).
$$

This is still gamma, but with 2α taking the place of α.