Risk & Asset Allocation
Homework for Week 1

John Dodson

September 3, 2014

Problems

Solutions to these problems are due at the beginning of the next session, which is 5:30 PM on Wednesday, September 10.

1. Post a document to the NetFiles dropbox (do not e-mail) telling me about your background and professional goals, your interests, what you expect to learn from this module, and how you think about the program so far. **(1 point)**

We saw that the entropy of a unit-variance normal random variable is $\log\sqrt{2\pi e} \approx 1.42$. The normal has "normal" tails. In this problem, let's look at the entropy of random variables with more extreme tails.

A uniform random variable $\mathcal{U}(\theta)$ for parameter $\theta > 0$ has a sample space $(0, \theta)$. The event space consists of the Borel sets on $(0, \theta)$, and the probability of each is the normalized Lebesgue measure; e.g., $P(a, b) = \frac{b-a}{\theta}$ for $0 < a < b < \theta$.

2. What value of the parameter, $\theta^\star$, corresponds to a unit variance? **(2 points)**

3. What is the entropy of the unit-variance $\mathcal{U}(\theta^\star)$? **(3 points)**

We can use the Chebychev inequality for a random variable X,

$$\mathrm{P}\left\{(X - \mathrm{E}\,X)^2 > x^2 \operatorname{var} X\right\} < \frac{1}{x^2} \quad \forall\, x > 1$$

to bound the distribution function for any symmetric random variable with a finite variance.

4. Based on this, what is an upper bound on the entropy for any symmetric, unit-variance random variable? **(4 points)**

Solutions

Firstly, note that since entropy transforms in a simple fashion under affine (linear plus a constant) transformations,

$$H_{aX+b} = \log|a| + H_X$$

for random variable X and constants a and b, it makes sense to limit our comparison to unit-dispersion random variables (unit-variance in this case).

Firstly, we need to determine the parameterization of a unit-variance uniform random variable. Let $U \sim \mathcal{U}(\theta)$. Since

$$\mathrm{E}\, U = \int_0^\theta u \frac{1}{\theta}\, du = \left. \frac{u^2}{2\theta} \right|_0^\theta = \tfrac{1}{2}\theta$$

and

$$\mathrm{E}\, U^2 = \int_0^\theta u^2 \frac{1}{\theta}\, du = \left. \frac{u^3}{3\theta} \right|_0^\theta = \tfrac{1}{3}\theta^2$$

the variance of U is $\operatorname{var} U = \frac{1}{3}\theta^2 - \left(\frac{1}{2}\theta\right)^2 = \frac{1}{12}\theta^2$. Therefore, for $\theta = \theta^\star = 2\sqrt{3}$, the variance is one.

In general, the entropy of a uniform random variable is

$$H_U = -\int_0^\theta \left(\log \frac{1}{\theta}\right) \frac{1}{\theta} du = \log \theta$$

so the entropy of a unit-variance uniform random variable is $\log 2\sqrt{3} \approx 1.24$, or about 12% less than that of the a unit-variance normal random variable.

Now, let's consider the Chebyshev inequality for symmetric mean-zero unit-variance random variables X.

$$\mathrm{P}\left\{|X| > x\right\} < \frac{1}{x^2} \quad \forall\, x > 1$$

Since $\mathrm{P}\{|X| > x\} = \mathrm{P}\{X > x\} + \mathrm{P}\{X < -x\} = 2\,\mathrm{P}\{X > x\} = 2\,(1 - \mathrm{P}\{X < x\})$,

$$F_X(x) > 1 - \frac{1}{2x^2} \quad \text{for } x > 1$$

A similar argument with $-x$ lead us to conclude that

$$F_X(x) < \frac{1}{2x^2} \quad \text{for } x < -1$$

Let Z be the random variable that attains these bounds. Since $F_Z(\cdot)$ must be non-decreasing,

$$F_Z(z) = \begin{cases} \frac{1}{2z^2} & z < -1 \\ \frac{1}{2} & -1 \le z \le 1 \\ 1 - \frac{1}{2z^2} & z > 1 \end{cases}$$

and, differentiating, we get the density

$$f_Z(z) = \begin{cases} -z^{-3} & z < -1 \\ 0 & -1 \le z \le 1 \\ z^{-3} & z > 1 \end{cases}$$

Note that Z is not actually unit-variance—in fact it has infinite variance—but any symmetric unit-variance random variable would have less mass in its tails than Z, and therefore lower entropy.

The entropy of Z is

$$H_Z = -2 \int_1^\infty \log\left(z^{-3}\right)\left(z^{-3}\right)\, dz = 6 \int_1^\infty z^{-3} \log z\, dz = \frac{3}{2}$$

Family	Entropy
uniform	$1.24 + \log \sigma$
Gauss	$1.42 + \log \sigma$
Chebychev	$< 1.50 + \log \sigma$

Table 1: Entropy of symmetric random variables with finite variance σ^2.

where the integral can be done by parts ($u = \log z$ etc.). This is only about 6% higher than the entropy of the unit-variance normal.

These results are summarized in Table 1.

One observation we can make is that the Gaussian random variable is actually fairly close to the upper bound. There is not much room unless we are prepared to abandon the requirement of a finite variance[1].

[1] *Update:* A student has since pointed out that the Gaussian is in fact provably the highest entropy random variable for a fixed variance.