Econ 506: Advanced Economic Statistics #1
Due on Monday 09/13/10

Hiroki Tsurumi
Wednesday September 1 2010

1. 10
2. 10
3. 6
4(i) 5
(ii) 5
5. 8
6. 10
~~54~~ 5?/60

> **1** In doing assignments throughout the term, collaboration among the students is encouraged. However, you must write your own answers. On the first page of your answer sheets you must write the collaborators:
>
> Assignment #1
> Name: Freddy Rojas Cama
>
> Collaborators: Keith Oberman, Rachel Madow, *etc.*
> or
> Antonio Cusato, Cesar tamayo
> Collaborators: None (if you worked on your own.)
>
> **2** You should use as many references as you wish, but you must give a list of references you used to answer each question. Put the list of references right after each question. If a reference is from the Internet you must give a full URL so that I can check the reference.
>
> **3** Always attach computer program(s) and output. However, you must report the computer output as a table, graph or equations in the text of your answer sheets.
>
> **4** As much as possible, word process your answers by LaTex, Scientific Tex, or other word processing software.
>
> **5** Staple together your answer sheets.
>
> **6** Throughout the term each question worths 10 points. "Sub-questions" (questions within a question) have equal weights.

1. Can you give an answer to **Exercise 1.14** on p.39 of **CB** that is different from the answer given in **ch1sol.pdf**?

2. Derive Bonferroni's inequality from Boole's inequality for $n = 3$.

3. Read relevant pages of Glenn Shafer and Vladmir Vovk (2005) "the original legacy of Kolmogorov's *Grundbegriffe*" to answer the following question:

> Question: What is the connection between Hilbert's sixth problem[1] and Kolmogorov's axiomatic treatment of probability?
>
> Give your answer in 50 words or less.

4. Read **Exercise 1.28** on p.40 of **CB** and their answer to the exercise. **CB**'s answer in **ch1sol.pdf** is sketchy. A more complete answer is given in "Calculus solver: Stirling's formula":

`http://www.sosmath.com/calculus/sequence/stirling/stirling.html`

The author of this proof says "Though the first integral is improper, it is easy to show that in fact it is convergent. Using the anti derivative of $\log(x)$ (being $x\log(x) - x$), we get...."

(i) Explain what the author means by "the first integral is improper, it is easy to show that in fact it is convergent."

(ii) Show

$$\int \log x \, \mathrm{d}x = x \log x - x.$$

Remark 1: In **Exercise 1.28** on p.40 **CB** says "a complete derivation of which is difficult. Instead, prove the easier fact,

$$\lim_{n\to\infty} \frac{n!}{n^{n+\frac{1}{2}}e^{-n}} = \text{constant."}$$

If you go to the history section of "*Stirling's approximation* in Wikipedia it becomes clear that Stirling's contribution is to show that this "constant" is $\sqrt{2\pi}$. The convergence of the left hand side is shown by de Moivre (1667–1754). Click "James Stirling" in the "Stirling's approximation" in Wikipedia you learn about James Stirling (1692–1770).

Remark 2: "Calculus Solver: Stirling's formula" gives a complete derivation of Stirling's formula. It is interesting to compare the **CB**'s derivation of

$$\lim_{n\to\infty} \frac{n!}{n^{n+\frac{1}{2}}e^{-n}} = \text{constant}''$$

[1] On p.82 of Reid, Constance (1970) *Hilbert*, Springer-Verlag, New York, Reid puts Hilbert's 6th problem as "to axiomatize those physical sciences in which mathematics plays an important role." Wikipedia writes Hilbert's 6th problem as "axiomatize all of physics." and says this problem is "unsolved."

to the derivation given in the Calculus Solver. (There are small typos in the **ch1sol.pdf**.

Remark 3: In **Exercise 1.28 CB** say "See **Exercise 5.35** for another derivation." This another derivation is close to the derivation of Stirling's formula by Walsh (1995). A complete reference to Walsh (1995) is given on slide p.31 of Pepe's lecture notes. Pepe gives Walsh's derivation on pp.17–20.

5. Write a GAUSS program to extend the table in slide 21 of Pepe's lecture notes to n=20, 30, 40, and 50. Attach your GAUSS program, and report the GAUSS output as a table.

6. Obtain the probability density functions (pdf's) of (a), (b), (c) and (d) of **Exercise 1.47** of **CB** and identify each distribution.

Advanced Economics Statistics

FALL 2010

First-Assignment Answer Sheet

1. This problem calls for the number of subsets that can be formed from elements of S. First, we should check that the proposed borel field (denoted by ß) includes all subsets of S and S itself. We define S_i as the collection of i-combinations of a set S; S_i is a subset of i distinct elements of S. Each set in S_i has different elements, formally $\cap_{i=0}^{n} S_i = \varnothing$. The number of elements in each S_i is defined by binomial formula $\binom{n}{i}$. We denote the number of elements of S_i as n_{S_i} and the k element in each subset S_i as $S_{i,k}$. First, we show that if we add a non-repeated element to S, the number of new elements in ß is $\sum_{i=0}^{n} n_{S_i}$ (the same number of elements we have before this new element belongs to S). Then, we will show that any proposed combination is already included in the ß, therefore ß contains all the combinations.

10

   We define S_{i+1}^{A} as a new set which takes the union of each element in S_i with a new set (A) of a single non-repeated element ($A = \{x \in N | x \cap S = \varnothing\}$). formally, $S_{i+1}^{A} = \left\{ S_{i+1,k}^{A} | S_{i+1,k}^{A} = (S_{i,k} \cup A) \cup S_{i+1} \right\}$, because A contains a non-repeated element in S, the number of elements of $\cup_{k=1}^{n_{S_i}} (S_{i,k} \cup A)$ is equal to the number of elements in S_i. The proposed borel field is $\left(\cup_{i=0}^{n} S_{i+1}^{A}\right) \cup \varnothing$ which has $\sum_{i=0}^{n} 2 \cdot n_{S_i}$ elements. In fact, the number of elements (or possible combinations) in this set is $\sum_{i=0}^{n+1} \binom{n+1}{i}$ where n is the number of elements of S. Thus, $2 \sum_{i=0}^{n} n_{S_i} = \sum_{i=0}^{n+1} \binom{n+1}{i}$. If we consider a initial set with $n-1$ elements then the borel field will contain : $2 \cdot \sum_{i=0}^{n-1} n_{S_i} = \sum_{i=0}^{n} n_{S_i}$. We can recursively do this process and finally we have: $2^n \sum_{i=0}^{1} n_{S_h} = \sum_{i=0}^{n+1} \binom{n+1}{i}$, being $\sum_{i=0}^{1} n_{S_i} = \binom{1}{0} + \binom{1}{1} = 2$.[1] Thus, $\sum_{i=0}^{n+1} \binom{n+1}{i} = 2^{n+1}$, and all elements in this set ß contains the i-combinations of elements in S. If we take any combination or element in $S_{j,l}$ which was constructed from S, by contruction, we have that $\left(\cup_{i=0}^{n} S_{i+1}^{A}\right) \cup \varnothing \cap S_{j,l}^{A} \neq \varnothing$ where $j = 0, ..., n+1$ and $l = 1, ..., n_{S_{j,l}^{A}}$. Thus, we show that these $\sum_{i=0}^{n} 2 \cdot n_{S_i}$ subsets or elements of S^A are all possible combinations of S^A and that ß is a borel field (definition 1.2.1 of Casella and Berger; 2002). □

2. This problem calls for deriving Bonferrony's inequality from Boole's inequality for $n = 3$. There are 3 sets which might be either joint or disjoint. We conduct this derivation in general way for 3 sets. For example, Figure 2A depicts a particular case when they are not pairwise disjoint.

[1] These 2 elements are related to values at the beginning and ending of each row in the pascal triangle. As we know the sum of each row i of this triangle gives us the number of i-combinations of a set.

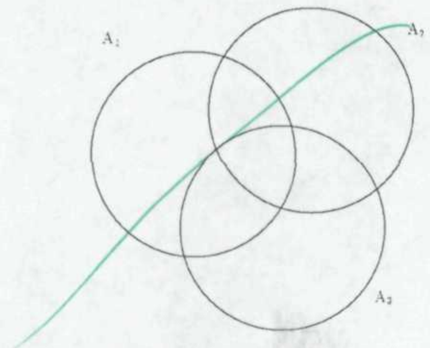

Figure 2A

Let's set up the derivation of Bonferroni's inequality when $n = 3$. Casella and Berger suggest to construct a disjoint collection of A_1^*, A_2^* and A_3^* which are a partition of $\cup_{i=1}^3 A_i$. So, $A_1^* = A_1, A_2^* = A_2 \cap A_1^c$ and $A_3^* = A_3 \cap (A_1 \cup A_2)^c$ .By construction we have that $A_2^* \subset A_2$ and $A_3^* \subset A_3$ so $\sum_{i=1}^3 P(A_i^*) \leq \sum_{i=1}^3 P(A_i)$. Then, it is true that $\sum_{i=1}^3 P(A_i^*) = P\left(\cup_{i=1}^3 A_i\right) = P\left(\cup_{i=1}^3 A_i^*\right)$. So, $P\left(\cup_{i=1}^3 A_i\right) \leq \sum_{i=1}^3 P(A_i)$. In order to get the final Boole's expression we need to use the following expression based on theorem 1.2.9 part (c) in Casella and Berger (2002); $P\left(\cup_{i=1}^3 A_i\right) \geq P\left(\left(\cap_{i=1}^3 A_i\right)^c\right)$ because $\left(\cap_{i=1}^3 A_i\right)^c \subset \cup_{i=1}^3 A_i$, the expression turns out to be an equality when the subset A_i form a partition of the sample space. So, it is true that $P\left(\left(\cap_{i=1}^3 A_i\right)^c\right) \leq \sum_{i=1}^3 P(A_i)$ (we call this expression 2.1) Considering Casella and Berger's expression $P\left(\cup_{i=1}^n A_i^c\right) \leq \sum_{i=1}^n P(A_i^c)$ [2], we have that expression 2.1 can rewrite as $P\left(\left(\cap_{i=1}^3 A_i\right)^c\right) \leq \sum_{i=1}^3 P(A_i^c)$, thus $1 - P\left(\cap_{i=1}^3 A_i\right) \leq 3 - \sum_{i=1}^3 P(A_i)$, then re-arranging terms we have $P\left(\cap_{i=1}^3 A_i\right) \geq \sum_{i=1}^3 P(A_i) - 2$. $\square$

Correct!

This is sufficient

3. What is the connection between Hilbert's sixth problem and Kolmogorov's axiomatic treatment of probability?. The most celebrated call about clarification and rigurosity of probability calculus came from Hilbert's six problem. Kolmogorov made a response; Kolmogorov focused on having a metric in the presence of elements in subsets and he came up with measures which neither are additive nor that the class of subsets to which it assigns numbers is a field. Further, Kolmogorov defined independence for partitions and stated that the characteristic of independence is responsable for power results in mathematics. Thus, Kolmogorov gave an important boost to the treatment of probability, even though conditional results and theory of distributions were missing at that time.

more than 50 words

4. This question is related to the proof of Stirling's formula.

   (i) The question calls for the meaning of improper integral. According to `mathworld`, an improper integral is a definite integral that has either both limits are infinite or an integrand that approaches infinity at one or more points in the range of integration. In terms of the problem, one of the integrals is improper because the integrand has a vertical asymptote at zero point. In the way to derive the Stirling's formula, an

[2] See section 1.2 in page 13 of Casella and Berger (2002)

useful expression is:

$$\int_{n-1}^{n} \log(x) < \log(n) < \int_{n}^{n+1} \log(x)$$

above inequality show the distance of $\log n$ with its neighborhood, this relation is true because the log function increases in a strictly-positive interval. Taking in consideration the sum of factorial n, we have

$$\int_{0}^{n} \log(x) < \log(n!) < \int_{1}^{n+1} \log(x) \tag{1}$$

5

thus, being n a finite number (this because we are interested in the factorial of a finite number) and the integrand of the lef-side integral goes to infinity when x approachs to zero (the integrand is not bounded). Thus, the first integral is improper because the integrand has a vertical asymptote at zero point. Stirling's procedure solves this problem using limits; if the limit exists we say that -in fact- the integral converges. First, we replace the lower limit of integral by a finite value called w, then we solve the "*indefinite*" integral using the antiderivative of the integrand:

$$\int_{w}^{n} \log(x) = x\log x - x|_{w}^{n} \tag{2}$$

we apply limits to (2)

$$\lim_{w \longrightarrow 0^+} f(n) - w\log w - w$$

being $f(n) = n\log n - n$, this expression is defined as long as $n > 0$. then,

$$\lim_{w \longrightarrow 0^+} w\log w - w$$

and above also can be expressed as

$$\lim_{w \longrightarrow 0^+} \frac{\log w}{w^{-1}} - \lim_{w \longrightarrow 0^+} w$$

applying L'hopital rule

$$\begin{aligned} \lim_{w \longrightarrow 0^+} \frac{\log w}{w^{-1}} &= \lim_{w \longrightarrow 0^+} -\frac{w^{-1}}{w^{-2}} \\ &= \lim_{w \longrightarrow 0^+} -w \\ &= 0 \end{aligned}$$

Thus we have that the left-hand integral in expression (1) is convergent

$$\int_{0}^{n} \log(x) = n\log n - n$$

(ii) The procedure for solving the integral is called integration by parts; we should follow the following steps:

$$\int \log x \ \mathrm{d}x = x \log x - x$$

we use the following useful expression:

$$d\left(uv\right) = du \cdot v + u \cdot dv$$

applying integrals

$$\int d\left(uv\right) = \int v \cdot du + \int u \cdot dv$$

5

now, we do that $x = u$ and $\log x = v$

$$\begin{aligned} x \cdot \log x &= \int \log x \cdot \mathrm{d}x + \int x \cdot \frac{1}{x} \mathrm{d}x \\ x \cdot \log x &= \int \log x \cdot \mathrm{d}x + x + C \\ x \cdot \log x - x &= \int \log x \cdot \mathrm{d}x + C \end{aligned}$$

where C is a constant. □

5. The GAUSS program for replicating the Pepe's lecture notes is shown in the appendix, this program is a user-friendly one and just you need to run and follows the instructions on the screen. We report the final table and convergence graph in this section:

```
Program to extend the table in slide 21 of Pepe's lecture notes to n=20, 30, 40 and 50.
Last Update September 7th 2010
By Freddy Rojas Cama
```

Table in slide 21 Pepe's lecture

n	n!	Stirling's formula	Relative error
20	2.43290201e+018	2.42278685e+018	0.00415765
30	2.65252860e+032	2.64517096e+032	0.00277382
40	8.15915283e+047	8.14217264e+047	0.00208112
50	3.04140932e+064	3.03634459e+064	0.00166526

Note: n! was approached using GAUSS software

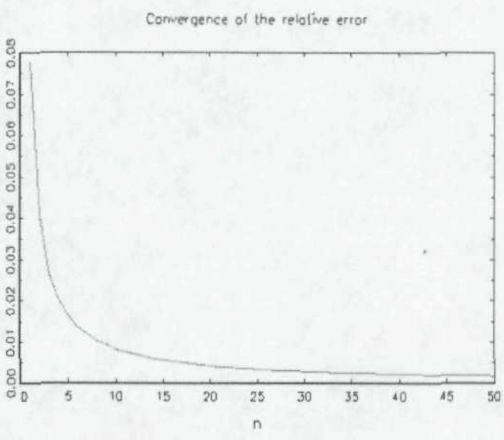

Figure 5.a

6. This exercise is related to exercise 1.47 of Casella and Berger (2002). First at all, we need to prove that the following functions are cdfs. According to Casella and Berger (2002), the theorem 1.5.3 states that a function is a cdf if and only if (a) $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1.$, b) $F(x)$ is a nondecreasing function of x and c) $F(x)$ is right-continuous. In each case $F(x)$ is a non-decreasing function, they are right continuos functions and $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$.

   (a) $F_X(x) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x), x \in (-\infty, \infty)$
   Considering that x has a support between $(-\infty, \infty)$ then we have that if $x \to -\infty$ then $\lim_{x\to-\infty} \tan^{-1}(x)$ - being $z = \arctan(x)$ - is equal to $-\frac{\pi}{2}$; so, $\lim_{x\to-\infty} \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x) = 0$. In the same way, if $x \to \infty$ then $\lim_{x\to\infty} \tan^{-1}(x)$ is equal to $\frac{\pi}{2}$; so, $\lim_{x\to\infty} \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x) = 1$. Then the probability density function (pdf) associated with this cumulative distribution is:

$$\begin{aligned} f_x &= \frac{\partial F}{\partial x} \\ f_x &= \frac{1}{\pi} \cdot \frac{1}{1+x^2} \end{aligned} \tag{3}$$

   being f_x a cauchy distribution with mode and entropy[3] parameters equal to 0 and 1 respectively.

   (b) $F_X(x) = (1 + \exp(-x))^{-1}, x \in (-\infty, \infty)$
   Considering that x has a support between $(-\infty, \infty)$ then we have that if $x \to -\infty$ then $\lim_{x\to-\infty} (1 + \exp(-x))^{-1} = 0$. In the same way, if $x \to \infty$ then $(1 + \exp(-x))^{-1} = 1$. Then the probability density function (pdf) associated with this cumulative distribution is:

$$f_x = \frac{\exp(-x)}{(1 + \exp(-x))^2}$$

   being a logistic distribution with mean and entropy parameter equal to 0 and 1 respectively.

   (c) $F_X(x) = e^{-e^{-x}}, x \in (-\infty, \infty)$, for sake of simplicity we transform the function using another monotonic function; being $\ln y = -e^{-x}$
   Considering that x has a support between $(-\infty, \infty)$ then we have that if $x \to -\infty$ then $\lim_{x\to-\infty} \ln y = -\infty$ and $y \to 0$ when the first occurs. In the same way, if $x \to \infty$ then $\lim_{x\to\infty} \ln y = 0$ and $y \to 1$ when the first occurs. Then the probability density function (pdf) associated with this cumulative distribution is:

$$\begin{aligned} g &= \frac{1}{y}\frac{\partial f}{\partial x} = e^{-x} \\ \frac{\partial f}{\partial x} &= e^{-x}e^{-e^{-x}} \end{aligned}$$

[3] In physics, entropy is defined as a measure of disorder in the system (see `mathworld` references).

being a *Gumbel* distribution with mode, *beta* and *gamma* parameter equal to 0, 1 and -1 respectively.[4] This distribution is a monotonic transformation of exponential distribution, being e^{-x} distribuited as exponential with *beta* parameter equal to 1, therefore $z = -ln(e^{-x})$ is *Gumbel*.

(d) $F_X(x) = \left(\frac{1-\epsilon}{1+e^{-x}}\right)$ if $x < 0$ and $F_X(x) = \epsilon + \left(\frac{1-\epsilon}{1+e^{-x}}\right)$ if $x \geq 0, x \in (-\infty, \infty)$
Considering that x has a support between $(-\infty, \infty)$ then we have that if $x \to -\infty$ then $\lim_{x\to-\infty} \left(\frac{1-\epsilon}{1+e^{-x}}\right) = 0$. In the same way, if $x \to \infty$ then $\lim_{x\to\infty} \epsilon + \left(\frac{1-\epsilon}{1+e^{-x}}\right) = 1$. Then the probability density function (pdf) associated with this cumulative distribution is:

Wrong

problem but ok

$$f_y = (1-\epsilon)\frac{e^{-x}}{(1+e^{-x})^2}$$

being a logistic distribution with mean and entropy parameter equal to 0 and 1 respectively.

References

[1] Casella, G and R. Berger. 2002. Statistical Inference. Second Edition, Duxbury Advanced Studies.

[2] Mendenhall and Scheaffer. 1973. Mathematical Statistics with Applications. Duxbury Press. North Scituate, Massachusetts.

[3] `Mathworld website. http://mathworld.wolfram.com/`

[4] According to Berger and Casella (2002)'s appendix if $\frac{1}{\beta}e^{-x\beta}$ is exponential, then $y = -\gamma \ln(\frac{1}{\beta}e^{-x\beta})$ is *Gumbel*.

1 Appendix

1.1 Gauss Program

```
/**********************************************/
    /**********************************************/
    /* Advanced Economic Statistics*/
    /* Assignment 1 */
    /* by Freddy Rojas Cama */
    /**********************************************/
    /**********************************************/
    /* Housekeeping and format*/
    /**********************************************/
    new;
    cls;
    library pgraph;
    pqgwin many;
    format /m1 /rd 20,2;
    " "
    " "
    " "
    " Professor Tsurumi's first assignment - Economics Department of Rutgers University"
    " Program to extend the table in slide 21 of Pepe's lecture notes to n=20, 30, 40 and 50.";
    " Last Update September 7th 2010 ";
    " By Freddy Rojas Cama";
    /************************************************/
    /* Settings */
    /************************************************/
    cls;
    " ";
    " ";
    " ";
    " In order to get the table (slide 21 Pepe's lecture notes) you should declare the factorial";
    " which you want to begin the table";
    n=con(1,1);
    nn=n;
    cls;
    " ";
    " ";
    " ";
    "How many rows do you want in your table (slide 21 Pepe's lecture notes)";
    rw=con(1,1);
    cls;
    " ";
    " ";
```

```
" ";
"Interval?";
ival=con(1,1);
cls;
/***************************************/
/*Matrix where I will save the results*/
/***************************************/
mat_={};
i=1;
do while i lt (rw+1);
 n_f=(2*pi*n)^0.5*(n/exp(1))^n;
 n_f_gauss=n!;
 rel_e=abs(n_f/n_f_gauss-1);
 row_=i~n~n_f_gauss~n_f~rel_e;
 mat_=mat_|row_;

 n=n+ival;
i=i+1;
endo;
/*******************************/
/*Format*/
/*******************************/
declare string fmtr = {"%6.0f","%6.0f","%11.8e","%11.8e","%6.8lf"};
s_mat_ = ftostrC(mat_, fmtr);
cls;
" ";
" Program to extend the table in slide 21 of Pepe's lecture notes to n=20, 30, 40 and 50.";
" Last Update September 7th 2010 ";
" By Freddy Rojas Cama";
" ";
" ";
" Table in slide 21 Pepe's lecture";
" ";; "————————————————————";
" ";; " n n! Stirling's formula Relative error";
" ";; "————————————————————";
" ";; s_mat_[.,2:cols(s_mat_)];
" ";
" ";; "————————————————————";
" ";; "Note: n! was approached using GAUSS software";
/*****************************************************/
/*mapping the approximation function and relative error*/
/*****************************************************/

graphset;
 _pcolor= 9;
```

CPU

NOT a ~~user~~ friendly FOR Loop.

Avoid when possible

No loop? why so, GAUSS exists??

```
_pmcolor = {1, 8,8, 8, 8, 8, 8, 2, 15};
_plwidth=4;
_paxht=0.2;
_ptitlht = 0.15;
@ _plegstr = "Fred";@
@_plegctl = { 1 10.25 };@
v_n=seqa(1,1,nn+ival*(rw-1));
v_n_f=(2*pi.*v_n)^0.5.*(v_n./exp(1)).^v_n;
v_n_f_gauss=v_n!;
v_rel_e=abs((v_n_f-v_n_f_gauss)./v_n_f_gauss);
title ("Convergence of the relative error");
xlabel("n");
ylabel("");
XY(v_n,v_rel_e);
/* END */
```