Gerald's Column
by Gerald Fitton

In my article this month I shall concentrate on my new found hobby, using email and its associated time wasting activity, surfing the world wide web. However, first I must respond to those who, I think, have been prompted to write to me by Colin Singleton's Standard Deviation puzzle in last month's Archive.

Standard Deviation

There is a rigorous Mathematical proof (so it must be 'True' mustn't it?) which leads to the $(n - 1)$ divisor. The proof does not require the population to be distributed according to any particular frequency distribution. What it does require is that the population of events from which the sample is drawn is either 'Infinite' (in practical terms this means very large) or that the samples are drawn 'with replacement' (by which I mean that as you select the events which go to make up your sample, you put them back into the pool before selecting the next event – thus the same event can appear more than once within the same sample).

Since Colin has an optimum divisor which varies with the sample size – sometimes below $(n - 1)$ and sometimes above it – then he is doing something which can loosely be described as 'wrong'. I remember a long time ago setting up a BBC computer to do something similar to that which I suspect Colin has done. I got 'wrong answers' and initially I blamed the BBC's random number selector thinking that the numbers selected were not random but only pseudo random. Then I spotted the fact that my algorithm was sampling without replacement. I changed the algorithm and the $(n - 1)$ divisor turned out to be the right one.

So, Colin and I differ on this one and I can't help anyone find the formula for the shape of Colin's curve because I think it should be a constant divisor.

For those of you who want to look up the proof that $(n - 1)$ is the optimum factor the Keyword or phrase is The Bessell Correction. Friedrich Wilhelm Bessell (1813–1898) was the famous mathematician who discovered that the correction was necessary (ie dividing by $(n - 1)$ instead of n). Like all mathematicians of that era he indulged in much Practical Mathematics (doing sums) as well as theoretical maths and was worried by discrepancies which arose in some of his more practical exploits.

Before I help you solve Colin's puzzle you'll have to show me where Colin's method differs from the assumptions used by Bessell.

Getting on Email

How did I resist joining this time consuming hobby for so long? Since my email address appeared in Archive a couple of months ago I must have received several hundreds of emails. The volume far exceeds the number of letters I would have received in the same period of time. I am truly grateful.

I would like to express my special thanks to those who have sympathised with the sadness

I feel at retiring from a job I have loved so much. Yes! I am quite seriously heartbroken – but I'll get over it; there is no point in feeling sorry for myself for too long.

Fortunately most of my hundreds of emails are brief and of the "Welcome to the Internet" type so that I have been able to answer them quite quickly. Do all these people read Archive – they must do!

Some have been more lengthy; often these longer emails make jocular or serious comment about my recent articles in Archive. Others have been requests for assistance with a problem. It has taken me a little longer to respond to this type of email so I hope that those of you who've sent me such correspondence will have found the delay bearable. A few of you who are used to more rapid responses to your emails have sent me reminders or polite requests asking if I've received the emails you've sent me. Sorry for the delay.

I have many things to say about email but, in a column such as that which Paul allows me, it is perhaps most appropriate to spend a little time answering the letters I have received by the so called Snail Mail (ordinary letters sent via the Post Office) asking whether I recommend "getting onto email". My response is that there are advantages and disadvantages to email communication and these will vary from individual to individual. For my own purposes the advantages outweigh the disadvantages and I wonder how I've resisted its call for so long.

Speed

On the plus side the communications are generally fast. I find that I can respond to an email question much more quickly than I can with a letter.

Rather than asking people to send me a file on disc which illustrates the problem they wish to solve or which illustrates the point they wish to make now they can attach such a file to an email. After initial problems no longer do I have any serious problems untangling the various attached files which come my way. I can send a reply back to them together with attached files illustrating my points with much less trouble than it takes to create a suitable set of files on disc.

On the down side people who use email seem to expect instant replies. I like to take my time over replying to complex communications even to the extent of drafting a reply, sleeping on it, tweeking it up the following morning and only then sending it off in the post. To those of you who email me with complex problems I would say "Don't expect an answer from me with the same alacrity that you get from most of your other email".

Data compression

Some of the files I have received have been very long. Indeed I have received emails from well over a dozen correspondents who have attached files longer than 100 Kb. One correspondent attached 2 Mb worth of files to a short email asking for assistance.

The modem with which Paul has presented me (a Robotics Sportster Vi 28.8 model), when used in conjunction with my aging A540 has a peak download rate of about 2000 bytes per second. I think that it is the 19 600 baud rate of the A540 Archimedes serial interface

which is the limiting feature rather than the modem. I'm sure that most of the readers of my column, being numerate, will have worked out already that 2 Mb will take a minimum of 1000 seconds (about 17 minutes) to download. In practice the time taken is usually longer than a calculation based on 2000 bytes per second because the computer sending the information has other jobs to do. Now and again it pauses for a few seconds or, on rare occasions, for about a minute or so.

Just to see what would happen I pushed the 2 Mb of files through !ArcFS, a compression utility, and, to my utter amazement, the 2 Mb reduced to 21 Kb! Usually the compression ratio is much less than 100:1; but just think, 10 seconds instead of 17 minutes.

I sent his compressed files together with my response, which I illustrated with more long but compressed files, back to my correspondent explaining what I had done and how to do it etc. My correspondent was equally amazed and responded with something like "So that's what these compression utilities can be used for".

Sending me compressed files

If you use a certain system to compress the files you send me you must be sure that I'll be able to decompress them. The decompression utilities I have include !SparkPlug (circa May 1996 version), !Squash version 0.45, !CFS version 1.10 and !ArcFS 2 version 2.11. Personally I like ArcFS best and CFS least. If you have information which imply that these versions are out of date then please let me know how to upgrade them.

A few days ago I received a compressed file from a software house. The file was to be used by me to upgrade software I have registered with them. It looked like an ArcFS folder but I just couldn't decompress it with my ArcFS. There were no instructions in the accompanying email specifying the utility needed for decompression. In the end and in desperation I tried !SparkPlug and it worked.

Also I have received compressed files which I can not decompress with Spark but can with ArcFS. They both look the same from the outside.

Hence my request to you, if your files are long then please compress them if you can – but do tell me the compression utility you've used.

The disadvantages of email

For many of my correspondents the biggest non advantage for them will be that the people with whom they wish to correspond will not be on email.

Of course I'm old enough to remember when having a telephone was rarer than email is today. In my view the telephone is a mixed blessing. The problem I find with it is that it can be intrusive. All too often I settle down in my favourite armchair after a hard day's work resolving not to answer any 'phone call; then the wretched thing goes off and my resolve is weakened by the possibility that it is not a double glazing firm but an emergency from my son or his wife.

I think that the number of people on email and who use it in preference to a telephone call

for all but emergencies will increase over the next few years. At least I can choose for myself the time at which I go on line and receive my emails – with the telephone I can't do that. I can't remember where I read it but it would seem that some company is setting up some sort of system whereby you can send and receive emails without having a full size computer. The gadget was rather like a palm top with a built in modem and all worked from a couple of AA size batteries. The price of these things are bound to decrease as time goes by and the advantage over the insidious telephone is, to me at least, self evident.

From time to time Jill and I stay in a hotel either for a short holiday or for business reasons (generally a mixture of both). I notice more of them have pay 'phone sockets suitable for plugging in a lap top or palm top. Some of these pay 'phone points embody a modem so that all you have to do is to provide the lap top and communications software. I'm sure that email communications are in the Development stage of the Life Cycle and that its use will grow.

When we get to the point that enough of our friends and relations are on email it will become impolite, except in an emergency, to use the old fashioned telephone since ia 'phone call will be perceived to be so intrusive. "It's bad to talk!"

I don't begrudge the cost of the 'phone calls to collect my emails. Indeed most of the calls are short, less than a couple of minutes. Even the long ones of 20 minutes or so won't cost me more than I can afford – but I do find it most frustrating waiting 20 minutes to find out what is coming down the line.

No! I think that the most serious disadvantage is this. I am reading an email which demands a well thought out response and feel an urgency about getting onto the next one. I have to resist giving a quick but inadequate reply. So what do I do?

I think I have developed a suitable protocol now. What I do is to put the email and attached files into a Pending directory for "I'll get around to it later" correspondence. So why is this a disadvantage? Well, on its own it isn't – but I had a quick look at my disc space a week ago and found that in a very short period of time I had gobbled up over 20 Mb with emails and their attached files. Of this 20 Mb around half was Pending. There is no doubt that the combination of storing emails and downloading all sorts of interesting pages from web sites uses up disc space like nothing I have experienced before.

The costs of email

I have decided to dedicate the whole of one 270 Mb Syquest removable disc to my Internet activities. This goes against the grain a bit for me because I'm not sure that I can handle such a massive amount of information. When I was a teacher (big sigh!) one of the points I always tried to make in my Statistics for Decision Makers lessons was that decision makers don't want all the facts. What they want is a reduced version which tells a story that they are able to understand. A phrase I didn't invent but used a lot was Information Overload. What I am trying to do is to work out a meaningful directory structure and search engine for all this data so that I can find what I want when I need to without having to open hundreds of unwanted files in my search. Neither do I want to repeatedly visit the same web site just because I can't find where on the disc I put that vital piece of information.

Anyway, disc space is certainly one cost of getting onto the Internet which I haven't heard

others mention.

In addition to the cost of disc space another consideration is the cost of starting up. You will need a modem and some software. I am fortunate in that Paul has presented me with both of these as a handsome retirement present. I am indebted to his persistence in persuading me to get on the net. What did I do before I had email and net surfing capabilities?

The software I am using is the ANT suite. I found it a bit difficult to start with since the instinctive DeskTop clicks didn't do what I expected. Nevertheless, after only a few days, I got to grips with what was required and now problems are few and far between. I haven't checked the detail of the prices (because it was a present) but for under £250 you should be able to buy a suitable modem and software. I understand that Paul has some 14.4 (KiloBaud?) modems going cheaply. This lower rate of 14.4 will give you a peak download rate of about 1500 bytes per second compared with the 2000 bytes per second of my system. It is not often that my system gets up to its peak rate. I find that the 2000 bytes per second of my system is quite adequate. I think that most of you would be quite happy with one of Paul's cheap 14.4 modems.

Service provider's charges

I didn't look into this as carefully as I might because I took the advice of a canny Scotsman (thanks John) whose advice to me has always been sound in the past. Since then I have corresponded with others about their choice of service provider.

My son, who uses a Windows machine, advises strongly against Compuserve for a variety of reasons only two of which I will list here. The first is that with Compuserve you have to start paying on line charges by the minute after a certain limit. The second is that the particular node which he was allocated gave him a lot of grief in terms of slowness of operation. His windows system should have been capable of a peak download rate of over 5500 bytes per second but his average rate was usually nearer 250 bytes per second. His particular interest was to download pictures of motor cars which were special in some way. Graphics do take a long time to download at the best of times. He took about an hour to get one picture of a '54 Chevrolet from a site in America. That was the last straw. His bill from Compuserve for one month was £80. Then he gave up the Internet.

I subscribe to Demon. It costs about £14 to register. That is a one time only charge. Then it is £10 per month plus VAT. Unlike Compuserve, this £11.75 per month give unlimited access. Demon also give, free of charge, 5 Mb of web space (see below). By commercial standards this is a small web site but adequate for the needs of most individuals.

Other people with whom I have corresponded susbscribe to Argonet. From what I have been told they run a service which is similar in style, cost and efficiency to that of Demon. There is one difference between Argonet and Demon which may be of interest to some of you. With Demon both Jill and I can have separate mail boxes at the same Internet site. By this I mean that you can send email to me at gerald@abacusline.demon.co.uk and it will go into my mailbox. If you want to send an email to Jill then write to jill@abacusline.demon.co.uk and it will go into her mailbox. Furthermore I can create other mail boxes such as webmaster@abacusline.etc at no extra cost. Each of these mail boxes can have their own password, their own mailing list, their own folders for incoming

and outgoing email, etc. I think that it is harder to do this with Argonet but I expect that there is a way around it.

Summary of Costs

The running costs are about £11.75 a month paid to the service provider plus the costs of the extra 'phone calls. It may be worth mentioning for those who don't know that all the calls are charged at the local rate and not the national rate. You can put your email provider's 'phone number on your BT Friends and Family list – in their Internet leaflet BT encourage you to do this.

The start up costs will run out at less than £250 unless you decide to go for a top of the range modem. If you go for the top of the range you might run up a bill for £400 with very little extra benefit. On the other hand you might pay less than £200 for a bottom of the range modem; a 14.4 which will not slow you down to an unacceptable extent.

Although some will disagree I would add that you should set aside 200 Mb of disc space for your Internet activities. If this means buying extra disc space then add it to your start up cost. In my case I have two external Syquest removable drives so my extra cost is really only the cost of an extra cartridge.

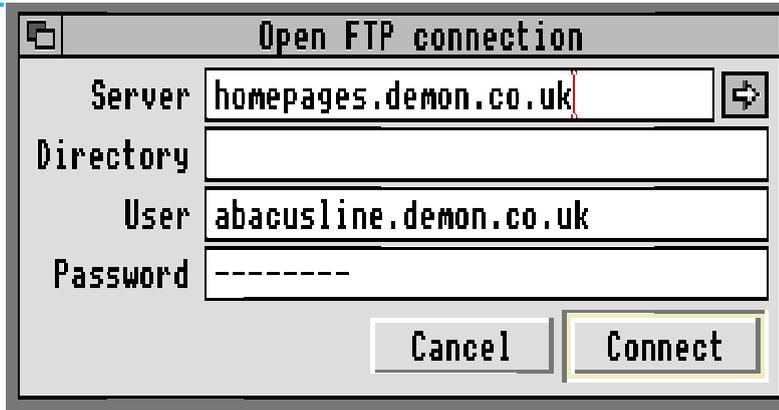## http://www.abacusline.demon.co.uk

I have now found out how to upload files to this site so, even as I write, there is something on the site. By the time you read this I should have quite a few of the free 5 Mb used up.

I have to confess that I am still experimenting and so whatever you find there today might be gone tomorrow! Nevertheless, here's an offer I would like you to take me up on. If you have a contribution which you think might be suitable for this site then send to to me as a file or directory attached to an email and I'll give it my consideration.

Web site creation

I had a look at the very sparse instructions which came from Demon about uploading files to my web site and had many problems filling in the dialogue box called "Open FTP connection" with the server address, Directory, User and Password. Everything I tried seemed not to work at first but, in the end I cracked it. The screenshot below shows the crucial dialogue. The Password is that provided to me by Demon when I paid my £14 to join. You will note that there is no http://www or ftp://www and no gerald@abacusline.etc.

The default supplied by Demon has the Server starting with "ftp://", the User as "Anonymous" and the Password as "gerald@abacusline.etc" – so most confusing.

```
┌──────────────────────────────────────────┐
│ ▣        Open FTP connection              │
├──────────────────────────────────────────┤
│   Server │homepages.demon.co.uk      │ ⇨ │
│ Directory│                           │   │
│     User │abacusline.demon.co.uk     │   │
│ Password │--------                   │   │
│                                          │
│              ┌────────┐ ┌─────────┐      │
│              │ Cancel │ │ Connect │      │
│              └────────┘ └─────────┘      │
└──────────────────────────────────────────┘
```

I read the articles provided by Demon, the articles in Archive and elsewhere and consulted with Paul about the software needed to create a web site. Matthew at NCS sent me a list of suitable software but remarked that he uses an unsophisticated text editor and a book rather than fancy software.

Anyway, I had a go using an unsophisticated text editor and learned the basics of HTML (HyperText Mark up Language). I uploaded a single file and then asked a couple of my email contacts to look at the page. To my great surprise I discovered that their Web Browsers produced similar but quite different layouts from that which I saw on my Fresco browser. I shouldn't have been surprised I suppose since my son had warned me that different Windows browsers produced different screen layouts.

HTML Portability

Long serving readers of this column will know of my interest in the portability of files from one package to another (eg from PipeDream to Schema) and across platforms and to different environments (eg from PipeDream on an Archimedes to Excel on a Windows machine). I have received letters from many correspondents with a wide variety of needs and from many with a great deal of experience in this field.

It occurred to me a long time ago that HTML might be a contender for a Portability Award. I never pursued this thought beyond a draft article for Archive which I never sent off. Now I have an opportunity to study the portability of HTML in great detail.

RTF Portability

Allow me to recall from long ago how one of my correspondents had studied the way in which RTF (Rich Text Format) ported between various packages and across platforms. He concluded that RTF was useful but not foolproof.

Paul sent me a utility which purported to convert RTF to HTML. I tried it with a measure of sucess. The main thing which had gone wrong was exactly the same as those of my RTF correspondent. One major problem was that **bold** was turning up everywhere where it wasn't wanted. When I looked at the HTML file in a text editor what I found was that many of the instructions such as the text here pair, which, in HTML language would make the words "text here" appear in bold type, were not nested properly within

other instructions such as <p>this is a paragraph</p>.

My first conclusion was that it was the RTF to HTML convertor which was producing things such as <p>some bold</p>text. The pair should be nested within the <p></p> pair and weren't. Further investigation now leads me to the conclusion that it is the Saving in RTF format from my word processor which may be the root cause of the problem. Anyway the method is unreliable.

HTMLEdit

I have used this software for only a few days but I have grown to like it. There are many useful utilities provided such as one which to all appearances looks like a filing system but allows the use of file names which are longer than the ten characters allowed by RISC OS. File names such as [PipeLine/html] would be truncated by RISC OS. The /html extension is highly desirable when uploading files to a web site. The method I used originally was to shorten the file name.

The editor and browser are superb and contain many keyboard short cuts. One feature of the editor which I found most useful is the syntax checker. This searches through your document for syntax errors and gives you a chance to correct them.

Web Browsers

Web browsers do more than read HTML files but one of their principal uses is as HTML interpreters. Different web browsers interpret the HTML commands differently even to the typeface and point size used for the body text. My son tells me that most Windows machine web browsers do not support italics in any legible way.

I have received correspondence from some who claim to know that even character sets can not be guaranteed beyond the first 128 basic ASCII codes – but I'm not convinced about this point since the HTLMEdit software seems to offer some sort of choice. On my web site I have included an address which includes accented characters, I have also included the £ symbol. If you visit my web site and find that these top bit set characters are producing rubbish for you then drop me an email and let me know what browser you're using.

Anyway, in my investigation of how to upload pages to my web site, I decided that one of the major considerations for the HTML commands used in my pages had to be that they had to produce something eminently readable with all the web browsers in common use on an Archimedes. Up to now I've discovered that the ANT browser I use, which is called Fresco, uses a smaller font size for headings than does many other browsers. The result is that my earlier attempts at a web page looked terrible for others because what I thought looked like well balanced headings spilt over onto the next line with some other browsers.

I must confess that I don't know how to tackle this problem of different interpretations HTML commands by different browsers except to ask those who visit my web site to let me know what it looks like. Please, if you send me a screenshot or a page saved as a Draw file then do compress the file and tell me what compression utility you've used.

Finally

If you able then then send me an email with attached files rather than send me a letter with an enclosed disc. If not then please send a disc, self addressed label and return postage if you want a problem solved or if you want to make a comment. My address is that of Abacus Training (see back inside cover).