STA 250/MTH 342 Intro to Mathematical Statistics
Lab Session 3 / Jan 26, 2015 / Handout

This session overviews graphical and numerical summarization of pdfs/pmfs and CDFs (cumulative distribution function) of scalar variables. Plotting the pdf/pmf/CDF curve is the obvious choice for a graphical summary. Numerical summaries can be based on various items, such as the expectation and variance under the pdf/pmf/CDF. We will discuss these keeping in mind applications to summarizing posterior pdfs/pmfs/CDFs in a Bayesian analysis. An additional challenge with such pdfs/pmfs is that they are often known up to a constant multiple (the normalizing constant that makes the curve a pdf or pmf).

See: `https://stat.duke.edu/courses/Spring15/sta250/labs/` for links to source code and data. Submit lab solutions via email to: `sta250@stat.duke.edu`. Any plots should be included in postscript form as attachments. The email subject must be "STA250 ..." with "..." replaced by your name.

1: Plotting pdf and CDF for $\mathsf{No}(\mu, \sigma^2)$.

Let f and F be the pdf and CDF of the normal distribution $\mathsf{No}(\mu, \sigma^2)$ respectively.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},$$
$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t.$$

Since both are defined on $(-\infty, \infty)$ which we cannot plot on a sheet of paper of finite size, we will only choose a small part of their domain. One usually chooses the interval $[\mu - 4\sigma, \mu + 4\sigma]$, partly because $F(\mu - 4\sigma) = 1 - F(\mu + 4\sigma) = 3.167124 \times 10^{-5}$ is very, very small.

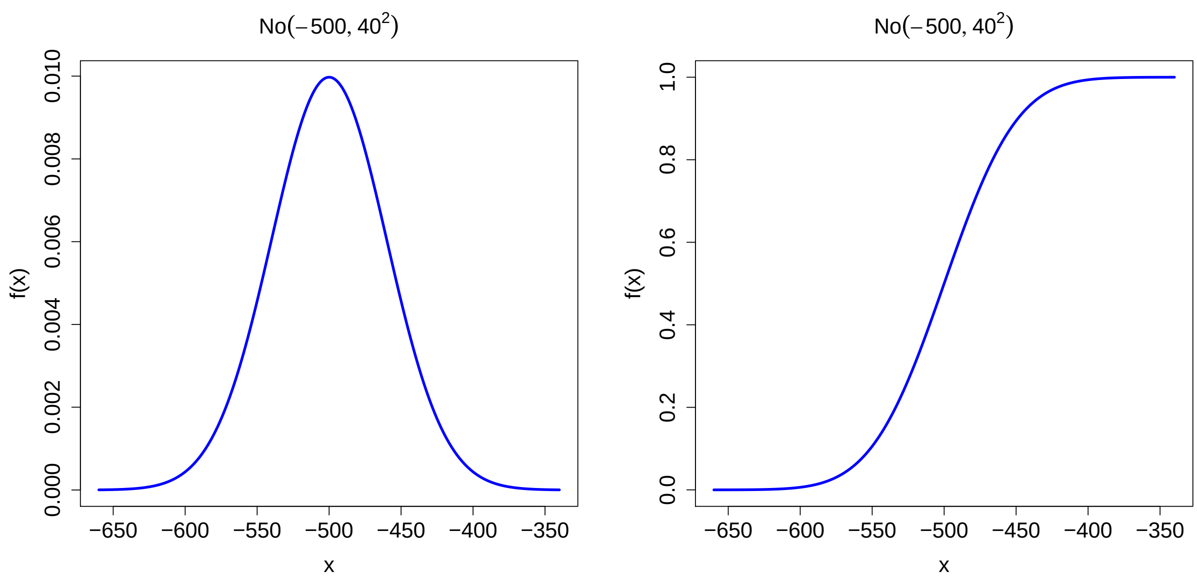

Figure 1: The pdf (left) and CDF (right) of $\mathsf{No}(-500, 40^2)$.

```
> mu <- -500
> sigma <- 40
> a <- seq(mu-4*sigma, mu+4*sigma, length.out = 1000)
> plot(a, dnorm(a, mean=mu, sd=sigma))#plot pdf
> plot(a, pnorm(a, mean=mu, sd=sigma))#plot CDF
```

One may wish to make a nice plot. The code that generated the pictures in Figure 1 is stored in "`gen.normal.plots.R`".

TASK 1 Modify the file "`gen.normal.plots.R`": change the color to red, make the lines thicker, correct the y-label in the second picture to $F(x)$, change the distribution to $\mathsf{No}(3000, 100^2)$, and generate your own pictures to submit as attachment. You may also do other modifications as you like.

2: Quantiles of $\mathsf{No}(\mu, \sigma^2)$. The R function "`qnorm`" solves for x the equation:

$$a = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \mathrm{d}t, \quad a \in [0,1];$$
$$x = \texttt{qnorm(}a\texttt{, mean = }\mu\texttt{, sd = }\sigma\texttt{)}.$$

TASK 2 Let F be the CDF of $\mathsf{No}(20, 5^2)$. Find an x so that $F(x) = 96.7\%$. Paste your command and result into the email.

3: Other Distributions. The **R** functions corresponding to other distributions are summarized below in Table 1.

Distribution	pdf/mdf	CDF	quantile	random var's
Binomial	dbinom	pbinom	qbinom	rbinom
Negative Binomial	dnbinom	pnbinom	qnbinom	rnbinom
Uniform	dunif	punif	qunif	runif
Poisson	dpois	ppois	qpois	rpois
Gamma	dgamma	pgamma	qgamma	rgamma
Normal	dnorm	pnorm	qnorm	rnorm
Hypergeometric	dhyper	phyper	qhyper	rhyper
Student t	dt	pt	qt	rt
Beta	dbeta	pbeta	qbeta	rbeta
χ^2	dchisq	pchisq	qchisq	rchisq
Exponential	dexp	pexp	qexp	rexp

Table 1: Summary of R functions about random distributions. For more information, read their documentation pages. For a full list, see `http://www.inside-r.org/r-doc/stats/Distributions`, for a more comprehensive list, see `http://cran.r-project.org/web/views/Distributions.html`

4: Plotting pmfs. The code to generate Figure 2 is saved in "`pmf.R`". In general, one need not memorize the commands in these plotting codes— just save the scripts and refer to them later when you need them.

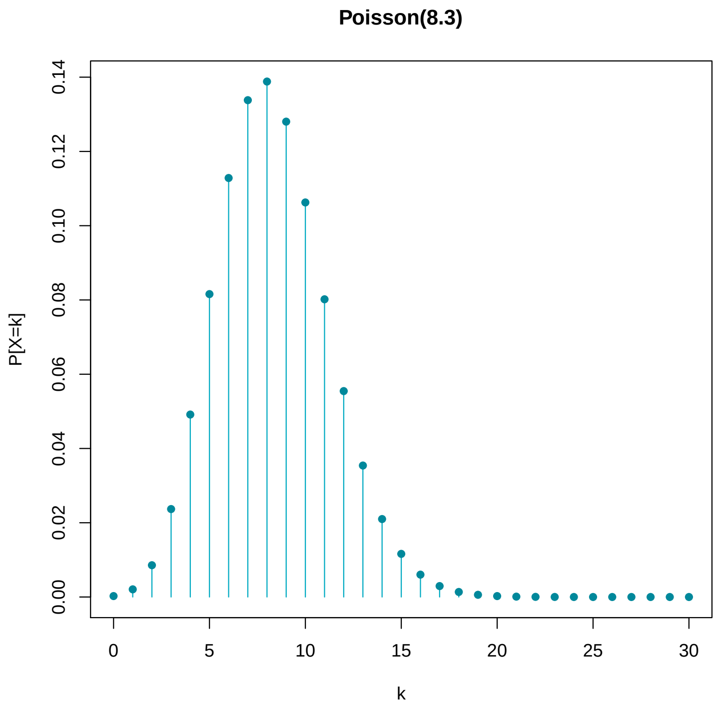

Figure 2: The pmf of $\mathsf{Po}(8.3)$.

TASK 3 Please modify the code in "`pmf.R`" to plot the pmf of the $\mathsf{Bi}(100, 0.40)$ distribution, for $k = 0, 1, \cdots, 25$.

~~END~~