# Statistics and Data Analysis, Fall 2015
# Pre-lecture Problems for Lecture 8: Statistical Estimation

Instructor: Ling-Chieh Kung
Department of Information Management
National Taiwan University

**Note.** DO NOT submit your answers. This problems are only for you to practice by yourselves. Doing these problems definitely help you understand course materials more. Of course, you are more than welcome to discuss these problems with the instructor or TA.

1. Recall that `NORMDIST(`$x$`, `$\mu$`, `$\sigma$`, 1)` can be used to find the left-tail probability of a normal distribution $\text{ND}(\mu, \sigma)$ with respect to a given threshold $x$. Moreover, `NORMINV(`$p$`, `$\mu$`, `$\sigma$`)` can be used to find the threshold such that the left-tail probability of a normal distribution $\text{ND}(\mu, \sigma)$ with respect to that threshold is $p$.

   (a) Let $X_i \sim \text{ND}(\mu, 4)$ be the population distribution and $\overline{X}$ be the sample mean of $X_i$ with 8 as the sample size. Show that $\overline{X} \sim \text{ND}(\mu, \sqrt{2})$ for any $\mu$.

   (b) Show that $\Pr(\overline{X} - \sqrt{2} \leq \mu \leq \overline{X} + \sqrt{2}) = 0.6827$ for any $\mu$.

   (c) Find $\Pr(\overline{X} - 2 \leq \mu \leq \overline{X} + 2)$ for any $\mu$. Compare your answer with the figure on page 16 of the slides.

   (d) Find `NORMINV(0.025, 10, SQRT(2))`. What does this value mean?

   (e) Find `NORMINV(0.025, 20, SQRT(2))`. What does this value mean?

   (f) Compare `10 - NORMINV(0.025, 10, SQRT(2))` and `20 - NORMINV(0.025, 20, SQRT(2))`. What does that mean?

   (g) Show that, for any $\mu$, if $\Pr(\overline{X} - b \leq \mu \leq \overline{X} + b) = 0.95$, then $b = 2.7718$.

   (h) Verify that $2.7718 = 1.96 \times \sqrt{2} =$ `-NORMINV(0.025, 0, 1)` $\times \sqrt{2}$. Explain this equality by reviewing the first bullet point of page 29 of the slides.

2. `TDIST(`$x$`, `$n-1$`, 1)` can be used to find $\Pr(T_{n-1} > x)$, the right-tail probability of a $t$ distribution with degree of freedom $n-1$ with respect to a given threshold $x$. `TDIST(`$x$`, `$n-1$`, 2)` can be used to find $\Pr(T_{n-1} > x) + \Pr(T_{n-1} < -x)$, the two-tail probability of a $t$ distribution with degree of freedom $n-1$ with respect to a given threshold $x$. Finally, `TINV(`$p$`, `$n-1$`)` can be used to find the critical $t^*$ value such that the two-tail probability of a $t$ distribution with degree of freedom $n-1$ with respect to $t^*$ is $p$ (i.e., $\Pr(T_{n-1} > t^*) + \Pr(T_{n-1} < -t^*) = p$).

   (a) Find `1 - TDIST(2.447, 6, 2)`. Convince yourself that the answer is the confidence level 95% specified on page 28 of the slides.

   (b) Find `1 - TDIST(2.447, 6, 1)`.

   (c) Find `TINV(0.05, 6)`. Convince yourself that the answer is the critical $t^*$ value for the 95% confidence level indicated on page 28 of the slides.

   (d) Suppose we want a 90% confidence level for the example on the slides, convince yourself that `TINV(0.1, 6)` is the critical $t^*$ value you need. Note that `TINV(0.1, 6)` is less than `TINV(0.05, 6)`. Why is that?

   (e) For a given sample size $n$ and a confidence level $1 - \alpha$, what are the parameters to input into `TINV()` to find the critical $t^*$ value?

   (f) Find `TDIST(TINV(0.1, 6), 6, 2)`. What does this value mean?