M4056 Lecture Notes. **September 3, 2010**

Sampling from a normal distribution

Recall:

- If X is $n(0,1)$, $M_X(t) = e^{t^2/2}$
- $M_{aX+b}(t) = e^{bt} M_X(at)$

If Y is $n(\mu, \sigma^2)$, then $Y = \sigma X + \mu$, where X is $n(0,1)$. Thus,

$$\text{if } Y \sim n(\mu, \sigma^2),\ M_Y(t) = e^{\mu\, t + \sigma^2\, t^2/2}.$$

Suppose $X_1, \ldots, X_n$ are any random variables. Then

$$M_{\overline{X}}(t) = E\, e^{t\overline{X}} = E\, e^{(t/n)(X_1 + \cdots + X_n)} = M_{X_1 + \cdots + X_n}(t/n).$$

We have seen previously that the *mgf* of distribution of a sum of independent random variables is the product of the *mfg*s of the summands. Hence:

$$\text{if } X_1, \ldots, X_n \text{ are } iid \sim X, \text{ then } M_{\overline{X}}(t) = [M_X(t/n)]^n.$$

(This appears in your book as **Theorem 5.2.7**.)

This enables us to calculate the distribution of the sum of *iid* normal variables.
If $X_1, \ldots, X_n$ are *iid* $n(\mu, \sigma^2)$, then

$$\begin{aligned} M_{\overline{X}}(t) &= [e^{\mu\,(t/n) + \sigma^2\,(t/n)^2/2}]^n \\ &= e^{n[\mu\, t/n + \sigma^2\,(t/n)^2/2]} \\ &= e^{\mu\, t + (\sigma^2/n)\,(t^2/2)} \end{aligned}$$

The last expression is the *mgf* of an $n(\mu, \sigma^2/n)$ random variable. This is consistent with the observations we made last time: the expected value and variance of $\overline{X}$ are as we have already seen they must be. But even more: the sample mean is itself normal, so these two numbers tell the whole story. We restate this important result:

Fact. If $X_1, \ldots, X_n$ are *iid* $n(\mu, \sigma^2)$, then $\overline{X}$ is $n(\mu, \sigma^2/n)$. (5.3.1.b)

Next, we turn to the distribution of the sample variance S^2 when samples of size n are drawn from an $n(\mu, \sigma^2)$ distribution.

Fact. $\overline{X}$ and S^2 are independent. (5.3.1.a)

Fact. $(n-1)S^2/\sigma^2$ has a χ^2_{n-1} (chi squared with $n-1$ degrees of freedom) distribution. (5.3.1.c)