MATH11400 Statistics 1 2010–11

Homepage http://www.stats.bris.ac.uk/%7Emapjg/Teach/Stats1/

Sampling distributions related to the Normal distribution

6.1 Revision of Moment Generating functions

(i) **Definition**

For a continuous random variable X the moment generation function $\mathcal{M}_X$ is defined by

$$\mathcal{M}_X(t) \equiv E(e^{tX}) = \int e^{tx} f_X(x)dx$$

(and by the corresponding sum in the discrete case), defined for whatever values of t the integral or sum is well defined. It uniquely determines the distribution, in the sense that two random variables with the same moment generating function have the same distribution.

(ii) **Some particular cases**

$$\begin{array}{llll} X \sim N(\mu, \sigma^2) & \iff & \mathcal{M}_X(t) = \exp\{\mu t + \sigma^2 t^2/2\} & t \in \mathbb{R} \\ X \sim \text{Exp}(\theta) & \iff & \mathcal{M}_X(t) = \theta/(\theta - t) & t < \theta \\ X \sim \text{Gamma}(\alpha, \beta) & \iff & \mathcal{M}_X(t) = \beta^\alpha/(\beta - t)^\alpha & t < \beta \end{array}$$

(iii) **Linear transformations**

If $Y = aX + b$ then $\mathcal{M}_Y(t) = E(e^{tY}) = E(e^{taX+tb}) = e^{tb}\mathcal{M}_X(ta)$

(iv) **Joint moment generating functions and Independence**

For two random variables X and Y, the joint moment generating function is defined as

$$\mathcal{M}_{X,Y}(s,t) \equiv \mathrm{E}(e^{sX+tY})$$

and similarly for more than 2 random variables.

The marginal moment generating functions for X and Y are given in terms of the joint moment generating function by

$$\begin{array}{rcccl} \mathcal{M}_X(s) & = & \mathrm{E}(e^{sX}) & = & \mathcal{M}_{X,Y}(s,0) \\ \mathcal{M}_Y(t) & = & \mathrm{E}(e^{tY}) & = & \mathcal{M}_{X,Y}(0,t) \end{array}$$

Two random variables X and Y are independent if and only if

$$\mathcal{M}_{X,Y}(s,t) = \mathcal{M}_X(s)\mathcal{M}_Y(t) = \mathcal{M}_{X,Y}(s,0)\mathcal{M}_{X,Y}(0,t).$$

(v) **Sums of independent random variables**

If $X_1, \ldots, X_n$ are independent and $Y = X_1 + X_2 + \cdots + X_n$, then

$$\mathcal{M}_Y(t) = E(e^{tY}) = E(e^{tX_1+\cdots+tX_n}) = E(e^{tX_1})E(e^{tX_2})\ldots E(e^{tX_n})$$

so

$$\mathcal{M}_Y(t) = \mathcal{M}_{X_1}(t)\mathcal{M}_{X_2}(t)\ldots\mathcal{M}_{X_n}(t)$$

If $X_1, \ldots, X_n$ are a random sample, i.e. they are all independent and all have the same distribution as a random variable X, then this simplifies to

$$\mathcal{M}_Y(t) = [\mathcal{M}_X(t)]^n$$

6.5 Proof of the independence of $\overline{X}$ and $\sum_{j=1}^{n}(X_j - \overline{X})^2$.

The proof is made up of the following steps:

(i) the definition gives an expression for the the joint moment generating function of $\overline{X}, X_1 - \overline{X}, X_2 - \overline{X}, \ldots, X_n - \overline{X}$ as the expected value of a complicated function of the variables;

(ii) simple manipulation reduces this complicated expression to a simple product of terms of the form $\exp\{a_j X_j\}$;

(iii) since the X_j are independent the expectation of this product is just the product of the expectations and each individual expectation is just the (known) moment generating function of X_j evaluated at a_j, giving a simple explicit expression for the joint mgf;

(iv) finally from the joint mgf we can calculate the marginal mgf for $\overline{X}$ and the (marginal) joint mgf for $X_1 - \overline{X}, X_2 - \overline{X}, \ldots, X_n - \overline{X}$ and note that the joint mgf is the product of these marginal mgfs.

Thus, by analogy with the result in the notes for two random variables, $\overline{X}$ is independent of the set of random variables $X_1 - \overline{X}, X_2 - \overline{X}, \ldots, X_n - \overline{X}$ and hence of $\sum_{j=1}^{n}(X_j - \overline{X})^2$.

Detailed Proof

(i) Let $\overline{X}$ denote $(X_1 + \cdots + X_n)/n$, let $\overline{s}$ denote $(s_1 + s_2 + \cdots + s_n)/n$ (so $\sum_{j=1}^{n}(s_j - \overline{s}) = 0$), and let $\mathcal{M}(t, s_1, \ldots, s_n)$ denote the joint moment generating function of the $n+1$ random variables $\overline{X}, X_1 - \overline{X}, X_2 - \overline{X}, \ldots, X_n - \overline{X}$.

Then by definition

$$\mathcal{M}(t, s_1, \ldots, s_n) = E(\exp\{t\overline{X} + s_1(X_1 - \overline{X}) + s_2(X_2 - \overline{X}) + \cdots + s_n(X_n - \overline{X})\})$$

(ii) Now rearranging the terms in the curly brackets gives

$$\begin{aligned} & t\overline{X} + s_1(X_1 - \overline{X}) + s_2(X_2 - \overline{X}) + \cdots + s_n(X_n - \overline{X}) \\ &= (t - s_1 - s_2 - \cdots - s_n)\overline{X} + s_1X_1 + s_2X_2 + \cdots + s_nX_n \\ &= a_1X_1 + \cdots + a_nX_n \qquad \text{for} \quad a_j = \frac{t - \sum_i s_i}{n} + s_j = \frac{t}{n} + (s_j - \overline{s}). \end{aligned}$$

(iii) Thus

$$\begin{aligned} & \mathcal{M}(t, s_1, \ldots, s_n) \\ &= E(\exp\{t\overline{X} + s_1(X_1 - \overline{X}) + s_2(X_2 - \overline{X}) + \cdots + s_n(X_n - \overline{X})\}) \\ &= E(\exp\{a_1X_1 + \cdots + a_nX_n\}) \\ &= E(\exp\{a_1X_1\})\, E(\exp\{a_2X_2\}) \cdots E(\exp\{a_nX_n\}) \quad \text{since each } X_j \text{ is independent} \end{aligned}$$

So $\mathcal{M}(t, s_1, \ldots, s_n)$

$$\begin{aligned} &= \mathcal{M}_{X_1}(a_1)\, \mathcal{M}_{X_2}(a_2) \cdots \mathcal{M}_{X_n}(a_n) \\ &= \exp\{\mu a_1 + \frac{\sigma^2 a_1^2}{2}\} \exp\{\mu a_2 + \frac{\sigma^2 a_2^2}{2}\} \cdots \exp\{\mu a_n + \frac{\sigma^2 a_n^2}{2}\} \quad \text{since each } X_j \sim N(\mu, \sigma^2) \\ &= \exp\{\mu\, \Sigma_j a_j + \frac{\sigma^2\, \Sigma_j a_j^2}{2}\} = \exp\{\mu t + \sigma^2 t^2/2n + \sigma^2 \sum_{j=1}^{n}(s_j - \overline{s})^2/2\}. \end{aligned}$$

where the last equality follows from the facts that $\sum_{j=1}^{n} a_j = t$
and $\sum_{j=1}^{n} a_j^2 = t^2/n + \sum_{j=1}^{n}(s_j - \overline{s})^2 + 2t\sum_{j=1}^{n}(s_j - \overline{s})/n = t^2/n + \sum_{j=1}^{n}(s_j - \overline{s})^2$.

(iv) Hence $\mathcal{M}(t,0,\ldots,0) = \exp\{\mu t + \sigma^2 t^2/2n\}$
and $\mathcal{M}(0,s_1,\ldots,s_n) = \exp\{\sigma^2 \sum_{j=1}^n (s_j - \overline{s})^2/2\}$
giving $\mathcal{M}(t,s_1,\ldots,s_n) = \mathcal{M}(t,0,\ldots,0)\,\mathcal{M}(0,s_1,\ldots,s_n)$.

Thus $\overline{X}$ is independent of the random variables $(X_1 - \overline{X}, X_2 - \overline{X}, \ldots, X_n - \overline{X})$ and in particular $\overline{X}$ is independent of $\sum_{j=1}^n (X_j - \overline{X})^2$.

6.9(b) Proof of the distribution of $\sum_{j=1}^n (X_j - \overline{X})^2/\sigma^2$.

First note that for any random variables $X_j,\ j = 1,\ldots,n$, with $\overline{X} \equiv (X_1 + \cdots + X_n)/n$,

$$\begin{aligned}\sum_{j=1}^n X_j^2 &= \sum_{j=1}^n (X_j - \overline{X} + \overline{X})^2 \\ &= \sum_{j=1}^n (X_j - \overline{X})^2 + 2\overline{X}\sum_{j=1}^n (X_j - \overline{X}) + \sum_{j=1}^n \overline{X}^2\end{aligned}$$

where, by definition, $\sum_{j=1}^n (X_j - \overline{X}) = 0$

so $\sum_{j=1}^n X_j^2 = \sum_{j=1}^n (X_j - \overline{X})^2 + n\overline{X}^2$. (as often seen before!)

Now consider first the case where $\mu = 0$ and $\sigma^2 = 1$, so $X_j,\ j = 1,\ldots,n$, is a random sample from the $N(0,1)$ distribution. Put $W_3 \equiv \sum_{j=1}^n X_j^2,\ W_2 \equiv \sum_{j=1}^n (X_j - \overline{X})^2$ and $W_1 \equiv n\overline{X}^2$. Then from above, $W_3 = W_2 + W_1$.

$$\begin{aligned}\text{Thus}\quad \mathcal{M}_{W_3}(t) &= \mathcal{M}_{W_1+W_2}(t) \\ &= \mathcal{M}_{W_1}(t)\mathcal{M}_{W_2}(t),\ \text{as } W_1 \text{ and } W_2 \text{ are independent from 6.4,} \\ \text{so}\quad \mathcal{M}_{W_2}(t) &= \mathcal{M}_{W_3}(t)/\mathcal{M}_{W_1}(t). \\ \text{But}\quad W_1 &\sim \chi_1^2 \ \text{from 6.5, as } \sqrt{n}\overline{X} \sim N(0,1) \text{ from 6.3 since } \mu = 0 \text{ and } \sigma^2 = 1, \\ \text{and}\quad W_3 &\sim \chi_n^2 \ \text{from 6.8(a),} \\ \text{so}\quad \mathcal{M}_{W_2}(t) &= (1-2t)^{-n/2}/(1-2t)^{-1/2} \\ &= (1-2t)^{-(n-1)/2} \quad \text{which is the mgf for the } \chi_{n-1}^2 \text{ distribution.} \\ \text{Thus}\quad W_2 &\equiv \sum_{j=1}^n (X_j - \overline{X})^2/\sigma^2 \sim \chi_{n-1}^2.\end{aligned}$$

To extend the proof to the case of general μ and σ^2, set $Y_j \equiv (X_j - \mu)/\sigma,\ j = 1,\ldots,n$, and set $W_3 \equiv \sum_{j=1}^n Y_j^2 = \sum_{j=1}^n (X_j - \mu)^2/\sigma^2,\ W_2 \equiv \sum_{j=1}^n (Y_j - \overline{Y})^2 = \sum_{j=1}^n (X_j - \overline{X})^2/\sigma^2$ and $W_1 \equiv n\overline{Y}^2 = n(\overline{X} - \mu)^2/\sigma^2$. Then the proof proceeds exactly as above, since W_3 and W_1 are again independent with χ_n^2 and χ_1^2 distribution respectively, so again we have $W_2 \equiv \sum_{j=1}^n (Y_j - \overline{Y})^2 = \sum_{j=1}^n (X_j - \overline{X})^2/\sigma^2 \sim \chi_{n-1}^2$.

6.14 The F distribution

(i) Definition

Let U and V be independent random variables with $U \sim \chi^2_r$ and $V \sim \chi^2_s$, and let $W = (U/r)/(V/s)$. Then we say W has the F distribution with r and s degrees of freedom and we write $W \sim F_{r,s}$.

(ii) Notes

1. The density function is heavily skewed with a long right tail.
2. If $W \sim F_{r,s}$ then, directly from the definition, $1/W \sim F_{s,r}$.
3. Define the percentage point $F_{r,s;\alpha}$ as the value such that $P(W \geq F_{r,s;\alpha}) = \alpha$ when $W \sim F_{r,s}$.
4. From (1–3) it follows that the distribution of F is not symmetric, but that the lower percentage points for the $F_{r,s}$ distribution can be found from the upper percentage points for the $F_{s,r}$ distribution, using the relationship $F_{r,s;1-\alpha} = 1/F_{s,r;\alpha}$.
6. For $x > 0$, the probability density function $f(x)$ for $F_{r,s}$ is given (see Rice §6.2) by

$$f(x) = \frac{\Gamma((r+s)/2)}{\Gamma(r/2)\Gamma(s/2)} \left(\frac{r}{s}\right)^{r/2} x^{r/2-1} \left(1 + \frac{rx}{s}\right)^{-(r+s)/2} .$$

(iii) Applications

Let $X_1, \dots, X_m$ be a random sample of size m from the Normal distribution $N(\mu_X, \sigma^2_X)$, and, independently, let $Y_1, \dots, Y_n$ be a random sample of size n from the Normal distribution $N(\mu_Y, \sigma^2_Y)$.

Let $\hat{\sigma}^2_X = \sum(X_i - \overline{X})^2/(m-1)$ and let $\hat{\sigma}^2_Y = \sum(Y_j - \overline{Y})^2/(n-1)$ be the usual variance estimators for the X and Y populations. Then these are independent and, from your notes, have distributions $\sum(X_i - \overline{X})^2/\sigma^2_X \sim \chi^2_{m-1}$ and $\sum(Y_j - \overline{Y})^2/\sigma^2_Y \sim \chi^2_{n-1}$.

Let

$$T = \frac{\sum(X_i - \overline{X})^2/(m-1)}{\sum(Y_j - \overline{Y})^2/(n-1)},$$

then when $\sigma^2_X = \sigma^2_Y =$ (say) σ^2

$$T = \frac{\sum(X_i - \overline{X})^2/(m-1)}{\sum(Y_j - \overline{Y})^2/(n-1)} = \frac{\sum(X_i - \overline{X})^2/\sigma^2_X(m-1)}{\sum(Y_j - \overline{Y})^2/\sigma^2_Y(n-1)} = \frac{\chi^2_{m-1}/(m-1)}{\chi^2_{n-1}/(n-1)} \sim F_{m-1,n-1},$$

and in particular the distribution of T is independent of the unknown parameters μ_X, μ_Y and σ^2.

This approach forms a starting point for statistical inference about the equality of variances for the two populations and, more generally, for inference in variety of linear models including linear regression and analysis of variance.