## $\boxed{6.4}$ Sums of R.V.'s.

### Convolution formula

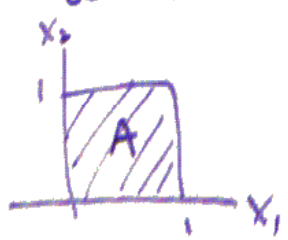
$S = X_1 + X_2 \Rightarrow f(x_1, x_2)$ joint pdf.

$$f_S(s) = \int_{-\infty}^{\infty} f(t, s-t)\, dt$$

Ex: Let $X_i \sim UNIF(0,1)$

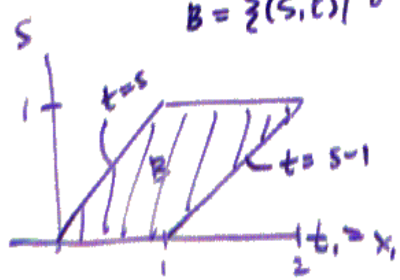
$X_1 \perp\!\!\!\perp X_2$

Let $S = X_1 + X_2$



$\Longrightarrow$

$B = \{(s,t) \mid 0 < t < s < t+1 < 2\}$



$$f_S(s) = \int_0^s dt = s, \quad 0 < s < 1$$

$$= \int_{s-1}^{1} dt = 2 - s, \quad 1 < s < 2$$

$$= 0, \quad \text{otherwise.}$$

### Moment generating function method

Thm: If $X_1, \dots, X_n$ are independent R.V. with MGF $M_{X_i}(t)$, then MGF of $Y = \sum_{i=1}^{n} X_i$ is

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

Proof: (Easy!)

$$M_Y(t) = E\left[e^{tY}\right] = E\left[e^{t\sum X_i}\right]$$

$$= E\left[e^{tX_1+tX_2+tX_3\cdots+tX_n}\right]$$

$$= E\left[e^{tX_1} e^{tX_2} e^{tX_3} \cdots e^{tX_n}\right]$$

$$= E\left[e^{tX_1}\right] E\left[e^{tX_2}\right] \cdots E\left[e^{tX_n}\right]$$

$$= \prod_{i=1}^{n} E\left[e^{tX_i}\right] = \prod_{i=1}^{n} M_{X_i}(t)$$

Note that if $X_i \sim iid$, then

iid = independent and identically distributed

$$M_Y(t) = \left[M_{X_1}(t)\right]^n$$

Ex: $X_i \sim BIN(n_i, p)$ $X_i$ is indep

Let $Y = \sum_{i=1}^{n} X_i$

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t) = \prod_{i=1}^{n} (pe^t + q)^{n_i} = (pe^t + q)^{\sum n_i}$$

$\Rightarrow Y \sim BIN(\sum n_i,$

Ex: $X_i \sim POI(\mu_i)$ $X_i$ is indep.

Let $Y = \sum_{i=1}^{n} X_i$

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(t) = \prod_{i=1}^{n} \exp\left[\mu_i (e^t - 1)\right] = \exp\left[\sum \mu_i (e^t - 1)\right]$$

$Y \sim POI(\sum \mu_i)$

Ex $\quad X_i \sim GAM(\beta, \alpha_i) \quad X_i \sim$ id $\quad$ d - independently distributed.

Let $Y = \sum_{i=1}^{n} X_i$

$$M_Y(t) = E(e^{tY}) = E(e^{t\sum X_i})$$

$$= E(e^{tX_1} e^{tX_2} \cdots e^{tX_n})$$

$$= E(e^{tX_1}) E(e^{tX_2}) \cdots E(e^{tX_n})$$

$$M_{X_1}(t) \; M_{X_2}(t) \cdots M_{X_n}(t)$$

$$\prod_{i=1}^{n} \left(\frac{1}{1-\beta t}\right)^{\alpha_i}$$

$$= \left(\frac{1}{1-\beta t}\right)^{\alpha_1} \left(\frac{1}{1-\beta t}\right)^{\alpha_2} \cdots \left(\frac{1}{1-\beta t}\right)^{\alpha_n}$$

$$= \left(\frac{1}{1-\beta t}\right)^{\alpha_1+\alpha_2+\cdots+\alpha_n}$$

$$= \left(\frac{1}{1-\beta t}\right)^{\sum \alpha_i}$$

So $Y \sim GAM(\beta, \sum \alpha_i)$

## 6.5 Order Statistics

Consider a random sample of data $X_1, \ldots, X_n$

Often it is useful to consider the "ordered" random sample, denoted by

$$X_{1:n}, X_{2:n}, \ldots, X_{n:n}$$

where $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$

The joint dist of $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ is not the same as joint density of the unordered variables

We will consider transformations that transform from $x_1, x_2, \cdots, x_n$ to $y_1, y_2, \cdots, y_n$ where

$$y_1 \le y_2 \le \cdots \le y_n$$

For example,

$y_1 = u_1(x_1, \cdots, x_n) = \min(x_1, \cdots, x_n)$

$y_2 = u_2(x_1, \cdots, x_n) =$ second smallest$(x_1, \ldots, x_n)$

$\vdots$

$y_i = u_i(x_1, \ldots, x_n) = i^{th}$ smallest$(x_1, \ldots, x_n)$

$\vdots$

$y_n = u_n(x_1, \ldots, x_n) = \max(x_1, \ldots, x_n)$

Thm If $X_1, X_2, \ldots, X_n$ is a RS from a population with continuous pdf $f(x)$, then the joint pdf of the order statistics $Y_1, \ldots, Y_n$ is

$$g(y_1, y_2, \ldots, y_n) = n!\, f(y_1) f(y_2) \cdots f(y_n)$$

Ex Suppose $X_1, X_2, X_3$ represent a RS from a pop with pdf

$$f(x) = 2x\, I_{(0,1)}(x)$$

$$g(y_1, y_2, y_3) = 3!\,(2y_1)(2y_2)(2y_3), \quad 0 < y_1 < y_2 < y_3 < 1$$

$$= 48\, y_1 y_2 y_3, \quad 0 < y_1 < y_2 < y_3 < 1$$

Marginal of $Y_1$?

$$g_1(y_1) = \int_{y_1}^{1} \int_{y_2}^{1} 48\, y_1 y_2 y_3 \, dy_3\, dy_2$$

= and then a miracle occurs.

$$= 6\, y_1 (1 - y_1^2)^2\, I_{(0,1)}(y_1)$$

= finish this problem on your own!

EX: X is cont $f(x) > 0$ on $a < x < b$
(a can be $-\infty$, and b may be $\infty$).

Let $n = 3$.

$$g_1 = \int_{y_1}^{b} \int_{y_2}^{b} 3!\, f(y_1) f(y_2) f(y_3)\, dy_3\, dy_2$$

$$= \int_{y_1}^{b} 3!\, f(y_1) f(y_2) \left[ \int_{y_2}^{b} f(y_3)\, dy_3 \right] dy_2$$

$$= \int_{y_1}^{b} 3!\, f(y_1) f(y_2) \left[ F(y_3) \right]_{y_2}^{b} dy_2$$

$$= \int_{y_1}^{b} 3!\, f(y_1) f(y_2) \left[ \underbrace{F(b)}_{1} - F(y_2) \right] dy_2$$

$$= 3!\, f(y_1) \int_{y_1}^{b} f(y_2) \left[ 1 - F(y_2) \right] dy_2$$

$u = 1 - F(y_2)$

$du = -f(y_2)\, dy_2$

$$= -3!\, f(y_1) \int_{1-F(y_1)}^{1-F(b)=0} u\, du = 3!\, f(y_1) \int_{0}^{1-F(y_1)} u\, du$$

$$= 3!\, f(y_1) \left. \frac{u^2}{2} \right]_0^{1-F(y_1)}$$

$$= \frac{3!\, f(y_1) \left[ 1 - F(y_1) \right]^2}{2}$$

$$= 3\, f(y_1) \left( 1 - F(y_1) \right)^2 I_{(a,b)}(y_1)$$

$$= 3\, (2y_1) \left( 1 - y_1^2 \right)^2 I_{(0,1)}(y_1)$$

Thm: Suppose $X_1, \dots, X_n$ denotes a RS of size $n$ from a cont. pdf $f(x)$, $f(x) > 0$ for $a < x < b$. Then the pdf of the $k^{th}$ order stat $Y_k$ is given by

$$g_k(y_k) = \frac{n!}{(k-1)!(n-k)!}\left[F(y_k)\right]^{k-1}\left[1-F(y_k)\right]^{n-k} f(y_k)\, I_{(a,b)}(y_k)$$

A mnemonic to memorize:

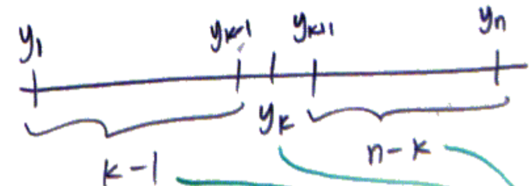
To have $Y_k = y_k$, one must have

$k-1$ observations less than $y_k$

one observation at $y_k$

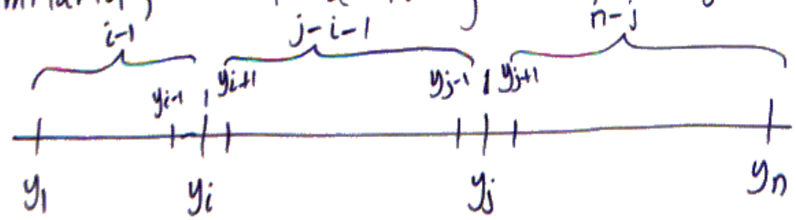
$n-k$ observations larger than $y_k$

There are $\dfrac{n!}{(k-1)!\,1!\,(n-k)!}$ possible orderings. Thus,




$$g_k(x_k) = \frac{n!}{(k-1)!\,1!\,(n-k)!}\left[F(y_k)\right]^{k-1} f(y_k) \left[1-F(y_k)\right]^{n-k}$$

Similarly, to find the joint pdf of $X_i$ & $X_j$, $i \neq j$.




$$g_{ij}(y_i, y_j) = \frac{n!}{(i-1)!\,1!\,(j-i-1)!\,1!\,(n-j)!}\left[F(y_i)\right]^{i-1} f(y_i)\left[F(y_j)-F(y_i)\right]^{j-i-1} f(y_j)\left[1-F(y_j)\right]^{n-j}, \text{ if } a<y_i<y_j<b$$

<u>Special order stats</u>:

smallest & largest.

sample range = $Y_n - Y_1$

Sample median $Y_k$, $k=\frac{n+1}{2}$ (n is odd)

min: $n\left[1-F(y_1)\right]^{n-1} f(y_1)$, $a<y_1<b$

max. $n\left[F(y_n)\right]^{n-1} f(y_n)$, $a<y_n<b$