ORDER STATISTICS :-

Suppose $x_1, x_2 \ldots, x_n$ are the elements of a random sample drawn from a univariate continious distribution having a density funⁿ $f(x)$. Let the values of x be rearrange in order of magnitude from least to greatest and be denoted by $x_{(1)}, x_{(2)}, \ldots x_{(n)}$. This rearrange n variates $x_{(n)}$ are called order statistics for the sample. Thus the rth order statistics of a sample of size n is simply the rth smallest obs in the sample and is denoted by $x_{(r)}$, r being the index of the order statistic for a conti. disⁿ P_r (prob) $[x_{(1)} < x_{(2)} < \ldots x_{(n)}] = 1$ The order statistics are not independent even though the original observation may be so examples of order statistics are extreme values. $x_{(1)}$ — the smallest and $x_{(n)}$ the longest, the m-th extreme $x_{(m)}$ the median and the ranges. The order statistics are widely used in SQC, in life testing exp, in climatology in determing non-parametric confidence internal and tolleranc regions.

Exact sampling dis^n of order statistics :-

Let $x_{(1)}, x_{(2)}, \cdots x_{(n)}$ be order statistics corresponding to n-indepent obs. x_i for $i=1,2,\cdots n$ drawn from a cont. dis^n with p.d.f. $f(x)$. Then the density of the order sample $x_{(1)} < x_{(2)} < \cdots < x_{(n)}$ for any specified permutation of sample is $f(x_{(1)}), f(x_{(2)}), \cdots f(x_{(n)})$ and there are in all $n!$ permutations of the n sample obs. which result in the same order statistics. i.e. the joint p.d.f. of $x_{(1)}, x_{(2)}, \cdots x_{(n)}$ is given by $g(x_{(1)}, x_{(2)}, \cdots x_{(n)})$

$$= n! \prod_{i=1}^{n} f(x_{(i)})$$

Sampling dis^n of the largest obs. $x_{(n)}$ in a random sample of n obs. drawn from a cont. population with p.d.f. $f(x)$:→

Let $x_{(1)} < x_{(2)} < \cdots x_{(n)}$ be n order statistic in a r.s. of size n, drawn from a cont. dis^n with p.d.f. $f(x)$. The allocation of n order statistics may be done in the following configuration.

$$-\infty \underbrace{\qquad\qquad}_{}\overset{(n-1)\,obs.}{} \quad x_{(n)} \quad x_{(n)} + dx_{(n)} \quad +\infty$$

such that $(n-1)$ obs. lie below $x_{(n)}$, 1 obs. bet^n $x_{(n)}$ and $x_{(n)} + dx_{(n)}$ lie and no observation lie above $x_{(n)} + dx_{(n)}$.

The prob. element of such allocation of 'n' obs. is by using multinomial prob. dis^n given by

$$dP(x_{(n)}) = \frac{n!}{(n-1)!\,1!\,0!}\left[\int_{-\infty}^{x_{(n)}} f(x)\,dx\right]^{n-1} \times \left[\int_{x_{(n)}}^{x_{(n)}+dx_{(n)}} f(x)\,dx\right]^{1} \times \left[\int_{x_{(n)}+dx_{(n)}}^{+\infty} f(x)\,dx\right]^{0}$$

$$= n\left[F(x_{(n)})\right]^{n-1} \times f(x_{(n)})\,dx_{(n)}$$

$$= n\left[F(x_{(n)})\right]^{n-1}\, dF(x_{(n)})$$

Sampling distribution of the smallest obs.:-

Let, $x_{(1)} < x_{(2)} < \cdots < x_{(n)}$ be 'n' order statistics in a r.s. of size n drawn ~~for~~ from a conti. dis^n with p.d.f. $f(x)$. We require to find the dist. of $x_{(1)}$. The allocation of 'n' order statistics may be done in the following ~~or~~ configuration.

$-\infty$ $|1|$ $(n-1)$ obs. $+\infty$

$x_{(1)}$ $x_{(1)} + dx_{(n)}$

$$\frac{n!}{0!\,1!\,(n-1)!}\left[\int_{-\infty}^{x_{(1)}} f(x)\,dx\right]^{0} \times \left[\int_{x_{(1)}}^{x_{(1)}+dx_{(1)}} f(x)\,dx\right] \times \left[\int_{x_{(1)}+dx_{(1)}}^{+\infty} f(x)\,dx\right]^{n-1}$$

such that no observation below $x_{(1)}$, I obs. in the interval $[x_{(1)}, x_{(1)}+dx_{(1)}]$ and $(n-1)$ obs. above $x_{(1)}+dx_{(1)}$.

The prob. element of such allocation is by using multinomial prob. dis^n given by

$$dP(x_{(1)}) = \frac{n!}{0!\,1!\,(n-1)!}\left[\int_{-\infty}^{x_{(1)}} f(x)\,dx\right]^0 \times \left[\int_{x_{(1)}}^{x_{(1)}+dx_{(1)}} f(x)\,dx\right]^1 \times \left[\int_{x_{(1)}+dx_{(1)}}^{\infty} f(x)\,dx\right]^{n-1}$$

$$= n\,f(x_{(1)})\,dx_1 \times [1 - F(x_{(1)})]^{n-1}$$

$$= n\,(1 - F(x_{(1)}))^{n-1}\,dF(x_{(1)})$$

This is the sampling dis^n of the smallest obs. in a r.s. of 'n' obs. drawn from a cont. dist. of $f(x)$

The Joint Distribution of rth and Sth ORDER Statistic $x_{(r)}$ and $x_{(s)}$ where $r<s$:—

The general allocation of 'n' order statistics for a r.s. of 'n' obs. drawn from a cont. dis^n is given by

$$\underset{-\infty}{}\;\underbrace{\quad}_{(r-1)}\;\underset{x_{(r)}}{}\;\underbrace{\quad}_{1}\;\underset{x_{(r)}+dx_{(r)}}{}\;\underbrace{\quad}_{(s-r-1)}\;\underset{x_{(s)}}{}\;\underbrace{\quad}_{1}\;\underset{x_{(s)}+dx_{(s)}}{}\;\underbrace{\quad}_{(n-s)}\;\underset{+\infty}{}$$

Now, applying multinomial prob. dist^n we have the Joint dist^n of $x_{(r)}$ and $x_{(s)}$

$$dP(x_{(r)}, x_{(s)}) = \frac{n!}{(r-1)!\,1!\,(s-r-1)!\,(n-s)!}$$

$$\left[\int_{-\infty}^{x_{(r)}} f(x)\,dx\right]^{r-1} \times \left[\int_{x_{(r)}}^{x_{(r)}+dx_{(r)}} f(x)\,dx\right]^{1}$$

$$\times \left[\int_{x_{(r)}+dx_{(r)}}^{x_{(s)}} f(x)\,dx\right]^{s-r-1} \times \left[\int_{x_{(s)}}^{x_{(s)}+dx_{(s)}} f(x)\,dx\right]^{1}$$

$$\times \left[\int_{x_{(s)}+dx_{(s)}}^{\infty} f(x)\,dx\right]^{n-s}$$

$$= \frac{n!}{(r-1)!\,(s-r-1)!\,(n-s)!} \times \left(F(x_{(r)})\right)^{r-1} dF(x_{(r)})$$

$$\times \left[F(x_{(s)}) - F(x_{(r)}+dx_{(r)})\right]^{s-r-1}$$

$$dF(x_{(s)}) \cdot \left[1 - F(x_{(s)})\right]^{n-s}$$

$$= \frac{n!}{(r-1)!\,(s-r-1)!\,(n-s)!} \left[F(x_{(r)})\right]^{r-1} \left[F(x_{(s)}) - F(x_{(r)})\right]^{s-r-1}$$

$$\left[1 - F(x_{(s)})\right]^{n-s} \cdot dF(x_{(r)})\, dF(x_{(s)})$$

This is the Joint dis^n of rth & sth order Statistics.

N.B.:- When $r=1$ & $s=n$ we get the Joint dis^n of smallest & largest order statistics.

$$dP(x_{(1)}, x_{(n)})$$

$$= n(n-1)\left[F(x_{(1)})\right]^0 \left[F(x_{(n)}) - F(x_{(1)})\right]$$

$$\left[1 - F(x_{(n)})\right]^0 \cdot dF(x_{(1)})\, dF(x_{(n)})$$

$$= n(n-1)\left[F(x_{(n)}) - F(x_{(1)})\right]^{n-2} dF(x_{(1)})$$

$$dF(x_{(n)})$$

Derive Joint dis^n of the smallest obs. $(x_{(1)})$ and largest obs. $(x_{(n)})$ of an order values can be made as follows :-

$$-\infty \quad |1| \quad (n-2) \quad |1| \quad +\infty$$

$$x_{(1)} \quad x_{(1)}+dx_{(1)} \qquad x_{(n)} \quad x_{(n)}+dx_{(n)}$$

'0' observation will fall below $x_{(1)}$, 1 observation in the interval $(x_{(1)}, x_{(1)}+dx_{(1)})$; $(n-2)$ obs in the interval $(x_{(1)}+dx_{(1)}, x_n)$, 1 obs in $(x_{(n)}, x_{(n)}+dx_{(n)})$ and '0' obs above $x_{(n)}+dx_{(n)}$

By applying multinomial prob. law we have the prob. disⁿ of such as as

$$dP(x_{(1)}, x_{(n)}) = \frac{n!}{0!\,1!\,(n-2)!\,1!\,0!}$$

$$\left[\int_{-\infty}^{x_{(1)}} f(x)\,dx\right]^0 \times \left[\int_{x_{(1)}}^{x_{(1)}+dx_{(1)}} f(x)\,dx\right]^1$$

$$\times \left[\int_{x_{(1)}+dx_{(1)}}^{x_{(n)}} f(x)\,dx\right]^{n-2} \times \left[\int_{x_{(n)}}^{x_{(n)}+dx_{(n)}} f(x)\,dx\right]^1$$

$$\times \left[\int_{x_{(n)}+dx_{(n)}}^{+\infty} f(x)\,dx\right]^0$$

$$= n(n-1)\, dF(x_{(1)}) \cdot \left[F(x_{(n)}) - F(x_{(1)})\right]^{n-2} dF[x_n] \cdot 1$$

This is the joint disⁿ of the sample obs. $x_{(1)}$ & $x_{(n)}$.

Distribution of the sample range: $\rightarrow w = x_{(n)} - x_{(1)}$

The joint disⁿ of smallest obs & largest obs. $x_{(1)}$ & $x_{(n)}$ is given by

$$dP(x_{(1)}, x_{(n)}) = n(n-1)\left[F(x_{(n)}) - F(x_{(1)})\right]^{n-2} dF(x_1) \cdot dF(x_n)$$

Put $w = x_{(n)} - x_{(1)}$

$u = x_{(1)}$

$\therefore x_{(1)} = u$

$x_{(n)} = w + u$

$$|J| = \frac{\partial(x_{(1)}, x_{(n)})}{\partial(u, w)} = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1$$

The Joint dis^n of u & w is given by

$$dP(u, w) = n(n-1)\left[F(w+u) - F(u)\right]^{n-2}$$

$$dF(u+w) - dF(u) \cdot 1$$

$$= n(n-1)\left[F(w+u) - F(u)\right]^{n-2}$$

$$f(u) \cdot f(u+w) : du \, dw$$

The dis^n of 'w' is obtained by integrating out 'u' over the ranges as

$$dP(w) = n(n-1)dw \int_u \left[F(w+u) - F(u)\right]^{n-2} f(u) \, f(u+w) du$$

Derivation of dist. of the rth order statistic $x_{(r)}$:→

Let, the rth order statistic in r.s. of size 'n' have the following configuration

$$\underset{-\infty}{\quad} \overset{(r-1)\text{ obs.}}{\rule{3cm}{0.4pt}} \underset{x_{(r)}}{|} \overset{1}{\rule{0.5cm}{0.4pt}} \underset{x_{(r)} + dx_{(r)}}{|} \overset{(n-r)}{\rule{3cm}{0.4pt}} \underset{+\infty}{\quad}$$

By applying multinomial prob. law we have the prob. dis^n of such allocation is given by

$$dP(x_{(r)}) = \frac{n!}{(r-1)!\,1!\,(n-r)!}\left[\int_{-\infty}^{x_{(r)}} f(x)\,dx\right]^{r-1} \times \left[\int_{x_{(r)}}^{x_{(r)}+dx_{(r)}} f(x)\,dx\right]^{1} \times \left[\int_{x_{(r)}+dx_{(r)}}^{+\infty} f(x)\,dx\right]^{n-r}$$

$$= \frac{n(n-1)(n-2)\cdots(n-r-1)}{(r-1)!}\left(F(x_{(r)})\right]^{r-1} \times f(x_{(r)})\,dx_{(r)}\,(1-F(x_{(r)}))]^{n-r}$$

$$= \left(\frac{n!}{(r-1)!\,(n-r)!}\,F(x_{(r)})^{r-1}\left[1-F(x_{(r)})\right]^{n-r} dF(x_{(r)})\right]$$

<u>Dist.^{n} of the sample median:-</u>

Let $x_{(1)}, x_{(2)}, \ldots x_{(n)}$ be the order statistics of a random sample of size n obs. drawn from a cont. population with p.d.f. $f(x)$.

Let us assume when 'n' is large $n = 2k+1$ then the $(k+1)$th obs. is the sample median having the following configuration

$$\underset{-\infty}{\quad} \overset{K}{\underline{\qquad\quad}}_{x_{(k+1)}} \quad |1| \quad \overset{K}{\underline{\qquad\quad}}_{(x_{(k+1)}+dx_{(k+1)})} \quad +\infty$$

K obs. lie below $x_{(k+1)}$, '1' obs i obs is $[x_{(k+1)}, x_{(k+1)} + dx_{(k+1)}]$ and K obs. lie above $x_{(k+1)} + dx_{(k+1)}$.

The prob. element of such allocation of $(2K+1)$ obs. is by using multinomial dist^n

$$dF(x_{(k+1)}) = \frac{(2K+1)!}{K!\,1!\,K!} \left(\int_{-\infty}^{x_{(k+1)}} f(x)\,dx \right)^K \left(\int_{x_{(k+1)}}^{x_{(k+1)}+dx_{(k+1)}} f(x)\,dx \right)^1 \left(\int_{x_{(k+1)}+dx_{(k+1)}}^{+\infty} f(x)\,dx \right)^K$$

$$= \frac{(2K+1)!}{(K!)^2} [F(x_{k+1})]^K \cdot dF(x_{(k+1)}) [1 - F(x_{(k+1)})]^K$$

writing $x_{(k+1)} = \tilde{x}$ we get the sampling dist^n of sample median as

$$dP(\tilde{x}) = \frac{(2K+1)!}{(K!)^2} F(\tilde{x})]^K dF(\tilde{x}) [1 - F(\tilde{x})]^K$$

Suppose a r.s. of 'n' obs is taken from a population with density func^n $f(x) = e^{-x}$, $x \geq 0$. Find the density func^n of the range & also show the mean value of 'w' is $1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{(n-1)}$

=> The Joint p.d.f. of the smallest obs. $x_{(1)}$ & the largest obs. $x_{(n)}$ in a random sample of size 'n' drawn from the given population is given by

$$dP(x_{(1)}, x_{(n)}) = n(n-1)\left[F(x_{(n)}) - F(x_{(1)})\right]^{n-2}$$

$$f(x_{(1)}) \cdot f(x_{(n)}) \cdot dx_{(1)} \cdot dx_{(n)}$$

where $F(x) = P(X \leq x)$

putting $w = x_{(n)} - x_{(1)}$

$U = x_{(1)}$

then $|J| = 1$

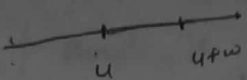
The Joint dist^n of 'w' & 'u' is given by $dP(w,u) = n(n-1)\left[F(u+w) - F(u)\right]^{n-2}$

$$f(u) \cdot f(u+w)\, du \cdot dw$$

The dist. of range 'w' is obtained by integrating out 'u' over the range as

$$dF(w) = n(n-1)\int_0^\infty [F(w+u) - F(u)]^{n-2} f(u)\, f(u+w)\, du$$

now $\int_u^{u+w} e^{-x}\, dx$



$$= -e^{-x}\Big]_u^{u+w}$$

$$= e^{-u}(e^{-w}-1)$$

$f(u) = e^{-u}$

$f(u+w) = e^{-u-w}$

$$dF(w) = n(n-1)\left\{e^{-u}(1-e^{-w})\right\}^{n-2} e^{-u-w}\, e^{-u}\, du$$

$$= n(n-1)e^{-w}(1-e^{-w})^{n-2}\int_0^\infty e^{-u(n-2)-2u}\, du$$

$$= n(n-1)e^{-w}(1-e^{-w})^{n-2}\int_0^\infty e^{-nu}\, du$$

$$= n(n-1)e^{-w}(1-e^{-w})^{n-2}\, dw \left.\frac{e^{-nu}}{-n}\right]_0^\infty$$

$$= n(n-1)(1-e^{-w})^{n-2}\,dw\left[-\frac{1}{n}\times 0+\frac{1}{n}\times 1\right]$$

$$=(n-1)(1-e^{-w})^{n-2}e^{-w}$$

$$E(w)=\int_0^\infty w\cdot f(w)\,dw$$

$$=\int_0^\infty w\cdot(n-1)(1-e^{-w})^{n-2}e^{-w}\,dw$$

$$=(n-1)\int_0^\infty w(1-e^{-w})^{n-2}e^{-w}\,dw$$

$$=(n-1)\int_0^1\left(z+\frac{z^2}{2}+\frac{z^3}{3}+\cdots\right)z^{n-2}\,dz$$

$$=(n-1)\int_0^1\left(z^{n-1}+\frac{z^n}{2}+\frac{z^{n+1}}{3}+\cdots\right)dz$$

$$=(n-1)\left[\frac{z^n}{n}+\frac{z^{n+1}}{2(n+1)}+\frac{z^{n+2}}{3(n+2)}+\cdots\right]_0^1$$

$$=(n-1)\left[\frac{1}{n}+\frac{1}{2(n+1)}+\frac{1}{3(n+2)}+\cdots\right]$$

$$=1-\frac{1}{n}+\frac{n+1-2}{2(n+1)}+\frac{n+2-3}{3(n+2)}+\frac{n+3-4}{4(n+3)}+\cdots$$

$$=1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}\cdot\text{———}\;+\frac{1}{(n-1)}.$$

$z=1-e^{-w}$

$1-z=e^{-w}$

$\log(1-z)=-w$

$\frac{1}{1-z}\,dz=dw$

$w=\log(1-z)^{-1}$

$=-\left\{-z-\frac{z^2}{2}-\frac{z^3}{3}-\right\}$