Stat 2911 Lecture Notes

Class 28, 2017

Uri Keich

Extrema and Order Statistics, Expectation of a continuous RV: Gamma, U(0,1) order statistics and its connection with probability plots, Cauchy, Normal, Function of a continuous RV (Rice 3.7, 4.1)

Extrema and Order Statistics

Let $X_1, \dots, X_n$ be iid RVs with CDF F.

$X_{(1)} \le X_{(2)} \le \dots \le X_{(n)}$ are the order statistics.

Claim. If F is a cont. function (so $\forall x$ $P(X=x)=0$)

then with prob. 1 $X_{(1)} < X_{(2)} \dots < X_{(n)}$.

Let $U = X_{(n)} = \max\{X_1, \dots, X_n\}$

$V = X_{(1)} = \min\{X_1, \dots, X_n\}$

$F_U(u) = [F(u)]^n$, $F_V(v) = 1 - P(V > v)$

If F has a density f,

$f_U(u) = n[F(u)]^{n-1} f(u)$, $f_V(v) = n\,[1-F(v)]^{n-1} f(v)$

What is the dist. of $X_{(k)}$?

$$F_{X_{(k)}}(x) = P(X_{(k)} \le x)$$
$$= P(\text{at least } k \text{ of } X_1,\dots,X_n \text{ are } \le x)$$
$$= \sum_{j=k}^{n} P(\underbrace{\text{exactly } j \text{ of } X_1,\dots,X_n \text{ are } \le x}_{|\{i : X_i \le x\}| = j})$$

Recall that $|\{i : X_i \le x\}| \sim$

$$\Rightarrow \qquad F_{X_{(k)}}(x) = \sum_{j=k}^{n} \binom{n}{j} [F(x)]^j [1-F(x)]^{n-j}$$

If F has a density then so does $X_{(k)}$ and by chain rule

$$f_{X_{(k)}}(x) = \sum_{j=k}^{n} \left\{ \binom{n}{j} j [F(x)]^{j-1} [1-F(x)]^{n-j} f(x) - \binom{n}{j}(n-j) [F(x)]^j [1-F(x)]^{n-j-1} f(x) \right\}.$$

Noting the identity $\binom{n}{j+1}(j+1) = \binom{n}{j}(n-j)$ this turns out to be a telescopic sum yielding

$$f_{X_{(k)}}(x) = k \binom{n}{k} [F(x)]^{k-1} [1-F(x)]^{n-k} f(x).$$

Alternatively, we can argue as follows.

Recall the law of total prob.: if A is an event and X is a discrete RV with $X(\Omega)=\{x_i\}$, then

$$P(A)=\sum_i P(A \mid X=x_i)\, P(X=x_i).$$

<u>Claim</u> For any event A and a cont. RV X

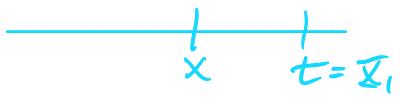
$$P(A)=\int_{-\infty}^{\infty} P(A \mid X=t)\, f_X(t)\, dt .$$

<u>Proof</u>. Beyond our current scope.

Take $A = A_x = \{(k)=1,\ X_{(k)} \le x\}$ and $X = X_1$ above, then the following makes sense intuitively:

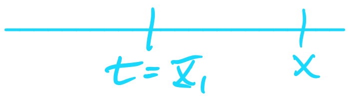
$$P(A \mid X_1=t) = P((k)=1,\ X_1 \le x \mid X_1=t) \quad \forall t \in \mathbb{R}$$

- for $t>x$

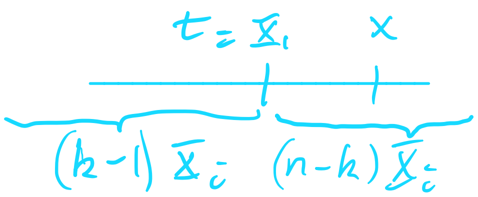


$$P(A \mid X_1=t) = 0$$

- for $t \le x$



$$P(A \mid X_1=t) = P((k)=1 \mid X_1=t)$$



$$P(A \mid X_1=t) = \binom{n-1}{k-1} [F(t)]^{k-1} [1-F(t)]^{n-k}$$

$$\Rightarrow P(A) = \int_{-\infty}^{x} \binom{n-1}{k-1} [F(t)]^{k-1} [1-F(t)]^{n-k} f(t)\, dt$$

Clearly, (the X_i are distinct with prob. 1)

$$P(X_{(k)} \le x) = \sum_{j=1}^{n}{}' P\left((k)=j,\ X_{(k)} \le x\right)$$

$$= n P\left((k)=1,\ X_{(k)} \le x\right)$$

symmetry

$$= n P(A).$$

Therefore,

$$P(X_{(k)} \le x) = \int_{-\infty}^{x} \underbrace{n \cdot \binom{n-1}{k-1} [F(t)]^{k-1} [1-F(t)]^{n-k} f(t)}\, dt$$

$$\Rightarrow f_{X_{(k)}}(x) = \uparrow$$

Recall: $n\binom{n-1}{k-1} = k \cdot \binom{n}{k}$. (expectation of binomial RV)

<u>Example.</u> $X_i \sim U(0,1)$

$$\Rightarrow F(x) = x \text{ and } f(x) = 1 \text{ for } x \in (0,1)$$

$$\Rightarrow f_{X_{(k)}}(x) = k\binom{n}{k} \underbrace{x^{k-1}}_{x^{a-1}} \underbrace{(1-x)^{n-k}}_{(1-x)^{b-1}} 1_{x \in (0,1)}$$

$$\Rightarrow X_{(k)} \sim \text{Beta}(\underset{\overset{\shortparallel}{a}}{k},\ \underset{\overset{\shortparallel}{b}}{n-k+1})$$

Expectation of a cont. RV

Recall: when studying discrete RVs we defined the expectation of a non-negative, or a non-positive, RV X with $X(\Omega)=\{x_i\}_i$ as

$$E(X)=\sum_i x_i\, p_X(x_i).$$

Analogously, for a cont. non-negative, or a non-positive, RV X we define its expectation as

$$E(X)=\int_{-\infty}^{\infty} x f_X(x)\,dx .$$

Note that if $X \geq 0$, $E(X)=\int_0^{\infty} x f_X(x)\,dx$ and $\int_{-\infty}^{0}$, if $X \leq 0$.

Examples.

1) $X \sim \Gamma(\alpha,\lambda)$ $\quad \alpha,\lambda>0$

$$E(X)=\int_0^{\infty} x \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}\,dx$$

$$=\frac{1}{\Gamma(\alpha)}\int_0^{\infty}(\lambda x)^\alpha e^{-\lambda x}\,dx$$

$t=\lambda x$

$$=\frac{1}{\Gamma(\alpha)}\int_0^{\infty} t^\alpha e^{-t}\,\frac{dt}{\lambda}$$

$$=\frac{\Gamma(\alpha+1)}{\Gamma(\alpha)}\,\frac{1}{\lambda}$$

$$=\frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)}\cdot\frac{1}{\lambda}=\frac{\alpha}{\lambda}.$$

$\Rightarrow$ If $X \sim \exp(\lambda)$ then $E(X)=$

2) $X_1, \dots, X_n$ are ind. $U(0,1)$ RVs

$$E(X_i) = \int_0^1 x \cdot 1 \, dx = \frac{1}{2} .$$

Recall that $X_{(k)} \sim \text{Beta}(k, n-k+1)$ with pdf

$$f_{X_{(k)}}(x) = k\binom{n}{k} x^{k-1}(1-x)^{n-k} \, 1_{x \in (0,1)} ,$$

$$\Rightarrow \quad E(X_{(k)}) = \int_0^1 x \, k\binom{n}{k} x^{k-1}(1-x)^{n-k} dx$$

$$= k\binom{n}{k} \int_0^1 x^k (1-x)^{n-k} dx$$

Recall: $\displaystyle\int_0^1 x^{a-1}(1-x)^{b-1} dx = \beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} .$

$$\Rightarrow \int_0^1 x^k (1-x)^{n-k} dx = \beta(k+1, n-k+1) = \frac{\Gamma(k+1)\Gamma(n-k+1)}{\Gamma(n+2)}$$

$$= \frac{k!(n-k)!}{(n+1)!} .$$

$$\Rightarrow \quad E(X_{(k)}) = k\binom{n}{k} \frac{k!(n-k)!}{(n+1)!}$$

$$= k \frac{n!}{(n+1)!}$$

$$= \frac{k}{n+1} .$$

Suppose F is a strictly $\uparrow$ and cont. CDF.

Recall that if Y is a RV with a CDF F then $U = F(Y)$ is a $U(0,1)$ RV.

Therefore, if $Y_1, \ldots, Y_n$ are a random sample from F then $U_i := F(Y_i)$ are a $U(0,1)$ sample.

Since $F \uparrow$, $U_{(k)} = F(Y_{(k)})$ so

$$E[F(Y_{(k)})] = E(U_{(k)})$$
$$= \frac{k}{n+1} \, .$$

Recall that in a prob. plot we draw

$$k/n+1 \quad \text{vs.} \quad F(y_{(k)}) \, .$$

It follows that <u>if</u> $Y_i \sim F$ then "on average" the prob. plot will go through the diagonal $y = x$ (in the unit square).

3) $\bar{X} \sim$ Cauchy $\quad f_{\bar{X}}(x) = \frac{1}{\pi}\frac{1}{1+x^2}$

$$E(\bar{X}) = \int_{-\infty}^{\infty} \underbrace{x \cdot \frac{1}{\pi}\frac{1}{1+x^2}}_{\text{odd}}\, dx =$$

The dist. is symmetric about 0.

SLLN: if $\bar{X}_i$ are iid with mean μ, then with $S_n = \sum_1^n \bar{X}_i$, $\frac{S_n}{n} \longrightarrow \mu$ with prob. 1.

This would imply that if $\bar{X}_i \sim$ Cauchy then $\frac{S_n}{n} \xrightarrow[n]{} 0$ with prob. 1. In particular, the variability of S_n/n diminishes to 0.

However, it can be shown that for $\bar{X}_i \sim$ Cauchy, S_n/n is itself a Cauchy RV!

The variability does not decrease and, in particular, S_n/n does <u>not</u> converge to 0. What is wrong?

Recall that in the discrete case, for a general RV $\bar{X}$, we could only define its expectation if it was in $L^1 = L^1(\Omega)$, which, in the discrete case, is: the space of RVs $\bar{X}: \Omega \longrightarrow \mathbb{R}$ for which

$$\sum_i |x_i| \, p_{\bar{X}}(x_i) < \infty.$$

For the same reason, for a cont. RV X that takes negative as well as positive values, we can only define its expectation if it is in L^1:

$$X \in L^1 \quad \text{if} \quad \int_{-\infty}^{\infty} |x| f_X(x)\, dx < \infty .$$

If $X \in L^1$ then its expectation is

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx \in \mathbb{R}$$

Going back to the Cauchy conundrum

$$\int_{-\infty}^{\infty} |x| \frac{1}{\pi} \frac{1}{1+x^2}\, dx > \int_{1}^{\infty} \frac{1}{\pi} \frac{x}{2x^2}\, dx$$

$$= \frac{1}{2\pi} \int_{1}^{\infty} \frac{1}{x}\, dx$$

$$= +\infty$$

$\Rightarrow X \notin L^1$ so $E(X)$ is <u>undefined</u>.

4) $Z \sim N(0,1)$ is $Z \in L^1$?

$$\int_{-\infty}^{\infty} |z| \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz = 2 \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} z e^{-z^2/2}\, dz = \sqrt{\frac{2}{\pi}} \left[-e^{-z^2/2} \right]_0^{\infty} = \sqrt{\frac{2}{\pi}}$$

$$\Rightarrow Z \in L^1 \quad \text{and} \quad E(Z) = \int_{-\infty}^{\infty} \underbrace{z \frac{1}{\sqrt{2\pi}} e^{-z^2/2}}_{\text{odd}}\, dz =$$

28.10

Recall: if X is a discrete RV with pmf p_X and range $X(\Omega)=\{x_i\}$, and if $Y=g(X)$ where $g:\mathbb{R}\to\mathbb{R}$, then

$$Y\in L^1 \iff \sum_i |g(x_i)|\,p_X(x_i)<\infty,$$

and if $Y\in L^1$ (or $Y\geq 0$, or $Y\leq 0$) then

$$E(Y)=\sum_i g(x_i)\,p_X(x_i).$$

If X is a cont. RV and $g:\mathbb{R}\to\mathbb{R}$ (measurable) then $Y=g(X)$ is not necessarily a cont. RV, nevertheless

$$Y\in L^1 \iff \int_{\mathbb{R}} |g(x)|\,f_X(x)\,dx<\infty,$$

and if $Y\in L^1$ (or $Y\geq 0$, or $Y\leq 0$) then

$$E(Y)=\int_{\mathbb{R}} g(x)\,f_X(x)\,dx.$$

<u>Proof</u> (assuming g is diff., strictly $\uparrow$ and onto $\mathbb{R}$).

$$Y=g(X) \implies f_Y(y)=f_X(g^{-1}(y))\,\underbrace{|g^{-1\prime}(y)|}_{=g^{-1\prime}(y)}$$

$$\implies \int_{-\infty}^{\infty}|y|\,f_Y(y)\,dy=\int_{-\infty}^{\infty}|y|\,f_X(g^{-1}(y))\,g^{-1\prime}(y)\,dy$$

$[y=g(x)\,;\,x=g^{-1}(y)]$ $$=\int_{-\infty}^{\infty}|g(x)|\,f_X(x)\,dx.$$

Therefore $Y\in L^1 \iff$ the last integral is $<\infty$. Repeating the argument without the $|\cdot|$ yields the second half.