Stat 110

Unit 8: Transformations

Ch. 8 in the text

Unit 8 Outline

- Change of Variables
- Convolutions
- Gamma Distribution
- Beta Distribution
- Beta-Gamma Connections
- Order Statistics (optional)

Change of Variables (1D)

- Let X be a continuous r.v. with PDF f_X, and let $Y = g(X)$, where g is differentiable and strictly increasing (or strictly decreasing). Then the PDF of Y is:

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right|$$

where $x = g^{-1}(y)$. The support of Y is all $g(x)$ with x in the support of X.

- This can be easily remembered based on the expression:

$$f_Y(y)dy = f_X(x)dx$$

- Note: if it's easier, you can find $|dy/dx|$ and take its reciprocal (but $f(y)$ should be expressed only in terms of y)

Change of Variables Example (1D)

- Let $X \sim N(\mu,\sigma^2)$ and let $Y = e^X$ (we saw in Unit 6 that the distribution of Y is Log-Normal). Find the PDF of Y.
- We are going to use the change of variables formula on the previous slide. So let's write some useful stuff down:

$y = e^x$ which means that: $x = \log(y)$ [that's base e of course]
So $dx/dy = d[\log(y)]/dy = 1/y$. Thus:

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right| = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\left|\frac{1}{y}\right| = \frac{1}{y\sqrt{2\pi\sigma^2}}e^{-\frac{(\log(y)-\mu)^2}{2\sigma^2}}$$

- And the support of y is $e^{-\infty} = 0$ to $e^{\infty} = \infty$.

Change of Variables (multiD)

- Let $\mathbf{X}$ be a continuous random vector with PDF $f_{\mathbf{X}}$, and let $\mathbf{Y} = g(\mathbf{X})$, where g is an invertible function from $\mathbb{R}^n$ to $\mathbb{R}^n$. Then the PDF of Y is:

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(\mathbf{x})\left|\frac{\partial \mathbf{x}}{\partial \mathbf{y}}\right|$$

The vertical bars mean "take the absolute value of the determinant of $\partial\mathbf{x}/\partial\mathbf{y}$".

- What is $\partial\mathbf{x}/\partial\mathbf{y}$?
- It's the *Jacobian matrix*, which is the matrix of all of the partial derivatives.

Jacobian Matrix

- The Jacobian matrix, $\partial \mathbf{x} / \partial \mathbf{y}$, is the matrix of all of the partial derivatives:

$$\frac{\partial \mathbf{x}}{\partial \mathbf{y}} = \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & & & \vdots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial y_n} \end{pmatrix}$$

- As in the 1D case,

$$\left|\frac{\partial \mathbf{x}}{\partial \mathbf{y}}\right| = \left|\frac{\partial \mathbf{y}}{\partial \mathbf{x}}\right|^{-1}$$

Change of Variables multiD Example

- Let $U \sim \text{Unif}(0, 2\pi)$ and $T \sim \text{Expo}(1)$, with T and U independent. Let $X = \sqrt{2T}\cos(U)$ and $Y = \sqrt{2T}\sin(U)$. Find the joint PDF of (X,Y). Are they independent? What are their marginal distributions?
- What is the joint PDF of U and T?

$$f_{T,U}(t,u) = \frac{1}{2\pi} e^{-t}, \text{ for } u \in (0,2\pi) \text{ and } t > 0.$$

- Viewing (X,Y) as a point in the place, what is the squared distance from the origin, in terms of U and T?

$$X^2 + Y^2 = 2T\cos^2(U) + 2T\sin^2(U) = 2T$$

Change of Variables multiD Example

- So $(\sqrt{2T}, U)$ is the polar coordinates expression of (X, Y).
- We can recover (T,U) from (X,Y), so the transformation is invertible and the change of variables formula applies.
- Let's start by calculating the Jacobian matrix:

$$\frac{\partial(x,y)}{\partial(t,u)} = \begin{pmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial u} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial u} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2t}}\cos(u) & -\sqrt{2t}\sin(u) \\ \frac{1}{\sqrt{2t}}\sin(u) & \sqrt{2t}\cos(u) \end{pmatrix}$$

  which has absolute determinant of $|\cos^2(u) + \sin^2(u)| = 1$.
- Note, we will need $\partial(t,u)/\partial(x,y)$ in the change of variables expression. So we need to invert the above (which is 1 ☺).

Change of Variables multiD Example

- So let's put everything together:

$$f_{X,Y}(x,y) = f_{T,U}(t,u)\left|\frac{\partial(t,u)}{\partial(x,y)}\right| = \frac{1}{2\pi} e^{-t} \cdot 1$$

$$= \frac{1}{2\pi} e^{-(x^2+y^2)/2} = \left(\frac{1}{\sqrt{2\pi}} e^{-x^2/2}\right)\left(\frac{1}{\sqrt{2\pi}} e^{-y^2/2}\right)$$

- For all real values of x and y.
- What does this factor into?
- Hole cow! We started with independent Uniform and Exponential, and ended up with two independent $N(0,1)$!
- This is called the Box-Muller method for generating Normal r.v.s. (It's very well known).

Change of Variables Outline

- Often times we may want to create a new random variable, $Z = g(X,Y)$, from some combination (not a simple addition) of two other random variables, X and Y (like for the Cauchy distribution in HW #9). Here's how you do it:
  1) Write down the joint PDF of (X,Y)
  2) Use the change of variables formula where $Z = g(X,Y)$, and simply define $W = X$ (or $W = Y$) to get the joint PDF of (Z, W)
  3) Integrate out the unwanted variable, W, and you are left with the PDF of the desired new r.v. Z.
- Sometimes step 3 is easy as the result may be two independent r.v.s.
- This is a very common approach in Stat 110 and 111.

Unit 8 Outline

- Change of Variables
- Convolutions
- Gamma Distribution
- Beta Distribution
- Beta-Gamma Connections
- Order Statistics (optional)

Convolutions

- If X and Y are independent discrete r.v.s, then the PMF of their sum $T = X + Y$ is:

$$P(T=t) = \sum_x P(Y=t-x)P(X=x)$$
$$= \sum_y P(X=t-y)P(Y=y)$$

- If X and Y are independent continuous r.v.s, then the PDF of their sum $T = X + Y$ is:

$$f_T(t) = \int_{-\infty}^{\infty} f_Y(t-x) f_X(x)dx$$
$$= \int_{-\infty}^{\infty} f_X(t-y) f_Y(y)dy$$

Convolutions Example #1: Exponentials

- Let i.i.d. $X, Y \sim \text{Expo}(\lambda)$. Find the distribution of $T = X + Y$.
- Since they are **independent** (a must) and we are trying to determine the distribution of the sum, we can use the convolution formula (continuous case here):
- First think: what are the bounds on $X = x$ here if we end up with $T = t$?

$$f_T(t) = \int_{-\infty}^{\infty} f_Y(t-x) f_X(x)dx = \int_0^t \left(\lambda e^{-\lambda(t-x)}\right)\left(\lambda e^{-\lambda x}\right)dx$$
$$= \lambda^2 \int_0^t e^{-\lambda t}dx = \lambda^2 e^{-\lambda t}\int_0^t dx = \lambda^2 t e^{-\lambda t}$$

- This is a known distribution which we'll see in a bit: $T \sim \text{Gamma}(2, \lambda)$

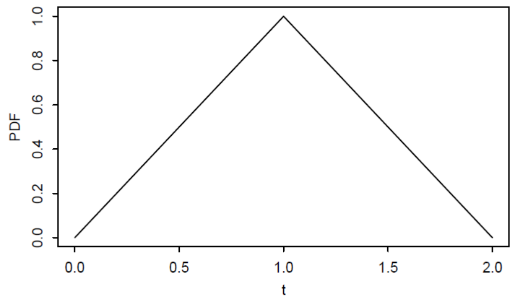
Convolutions Example #2: Uniform

- Let i.i.d. $X, Y \sim \text{Unif}(0,1)$. Find the distribution of $T = X + Y$.
- Again we can use the convolution formula:
- First think: what are the bounds on $X = x$ here if we end up with $T = t$? And what are the bounds on t?
  $0 < x < 1$ and $0 < t - x < 1 \rightarrow x > t - 1$ and $x < t$.
- So the lower bound on x is $\max(0, t - 1)$ and upper bound is $\min(1, t)$. Thus we break the integral into two parts (depends on how t compares to the value 1):

$$f_T(t) = \begin{cases} \int_0^t dx = t & \text{for } 0 < t \le 1 \\ \int_{t-1}^1 dx = 2 - t & \text{for } 0 < t \le 1 \end{cases}$$

- What's that plot look like? That's why it's called the triangle distribution!

Plot of Resulting Triangle Distribution



- What distribution does this resemble?
- The results of adding up 2 dice rolls! That's the discrete analogue to this problem: sum of i.i.d. $X,Y \sim \text{DUnif}(1,2,\ldots,6)$.

Unit 8 Outline

- Change of Variables
- Convolutions
- Gamma Distribution
- Beta Distribution
- Beta-Gamma Connections
- Order Statistics (optional)

Story of the Gamma Distribution

- The Gamma distribution is a generalization of the Exponential distribution. Instead of waiting for one success under memorylessness, a Gamma r.v. represents the total waiting time for multiple successes (a of them in fact).
- We already saw that if we are summing up two i.i.d. Expo(λ), then the result will be Gamma($2, \lambda$)
- If we sum up n i.i.d. Expo(λ), then the result will be Gamma($a = n, \lambda$)
- However in the Gamma distribution, a is not forced to be an integer (but has that nice relationship to the Exponential distribution if it is).
- Note, Gamma($1, \lambda$) is the same distribution of Expo(λ).

Gamma Distribution Definition

- A r.v. X has a ***Gamma distribution***, $X \sim \text{Gamma}(a, \lambda)$, with parameters a and λ ($a > 0$ and $\lambda > 0$) if its PDF is:

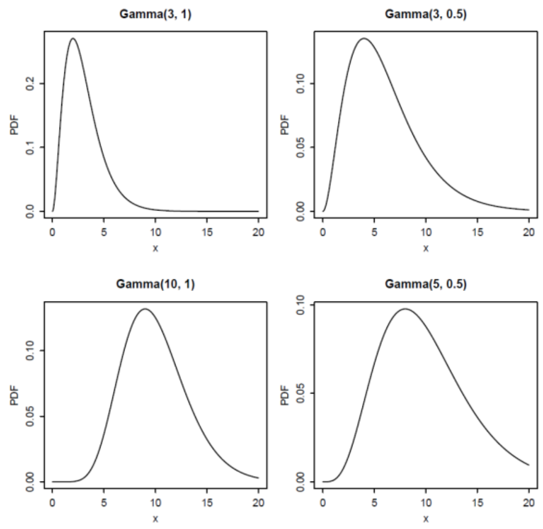
$$f(x) = \frac{1}{\Gamma(a)}(\lambda x)^a e^{-\lambda x}\frac{1}{x} = \frac{1}{\Gamma(a)}(\lambda x)^{a-1}\lambda e^{-\lambda x}$$

  for $x > 0$, where $\Gamma(a)$ is chosen to make the PDF integrate to 1.

- By definition, the normalizing constant, sometimes called the *Gamma function*, satisfies:

$$\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}dx$$

Plots of Gamma PDFs



That pesky Gamma function

- Recall, the normalizing constant, the *Gamma function*:

$$\Gamma(a) = \int_0^\infty x^{a-1}e^{-x}dx$$

- Here are some nice properties of it (just do integration by parts of the above integral to show the following hold):
  - $\Gamma(a + 1) = (a)\Gamma(a)$ for all $a > 0$.
  - $\Gamma(n) = (n-1)!$ if n is a positive integer.
- These properties makes things tenable, especially when calculating Expectations. We will use them along with **doing integrals by pattern (PDF) recognition**!

More Gamma Details

- Let $X \sim \text{Gamma}(a,1)$. Calculate $E(X)$ and $\text{Var}(X)$. Hint: use ***PDF recognition*** to calculate the integral.

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

$$= \frac{\Gamma(a+1)}{\Gamma(a)}\int_0^\infty \frac{1}{\Gamma(a+1)}(x)^{a+1}\frac{e^{-x}}{x}dx = \frac{a\Gamma(a)}{\Gamma(a)}\cdot 1 = a$$

$$E(X^2) = \int_0^\infty x^2\frac{1}{\Gamma(a)}x^a\frac{e^{-x}}{x}dx = \int_0^\infty \frac{a(a+1)}{\Gamma(a+2)}x^{a+2}\frac{e^{-x}}{x}dx = a(a+1)$$

$$\text{Var}(X) = E(X^2) - [E(X)]^2 = a(a+1) - a^2 = a$$

Gamma Mean and Variance

- Let $X \sim$ Gamma(a,1) and define $Y = X/\lambda$. Show that $Y \sim$ Gamma(a,λ), and calculate $E(Y)$ and Var(Y).
- Note: $y = x/\lambda$ means that $x = \lambda y$.

$$f_Y(y) = f_X(x)\left|\frac{dx}{dy}\right| = \frac{1}{\Gamma(a)} x^a \frac{e^{-x}}{x}\lambda$$

$$= \frac{1}{\Gamma(a)}(\lambda y)^a \frac{e^{-\lambda y}}{\lambda y}\lambda = \frac{1}{\Gamma(a)}(\lambda y)^a \frac{e^{-\lambda y}}{y}$$

- This is the PDF of a Gamma(a,λ) distribution.
- $E(Y) = E(X/\lambda) = a/\lambda$. $\mathrm{Var}(Y) = \mathrm{Var}(X/\lambda) = \mathrm{Var}(X)/\lambda^2 = a/\lambda^2$.

A sum of n i.i.d. Expo(λ) is Gamma(n,λ)

- Let i.i.d. $X_1,\ldots, X_n \sim$ Expo(λ). Show that $Y = X_1+\ldots+X_n \sim$ Gamma(n, λ) based on MGFs.
- Recall the MGF of an Expo(λ) is $\lambda/(\lambda-t)$. Thus the MGF of $Y = X_1+\ldots+X_n$ is $M_Y(t) = [\lambda/(\lambda-t)]^n$.
- The MGF of a Gamma(n, λ) is (good old PDF recognition):

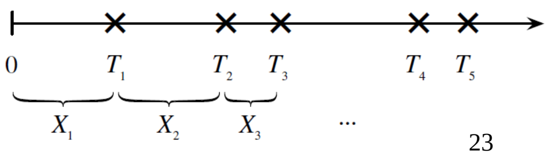
$$E(e^{tY}) = \int_0^\infty e^{tY}\frac{1}{\Gamma(n)}(\lambda y)^n \frac{e^{-\lambda y}}{y}dy$$

$$= \int_0^\infty \frac{1}{\Gamma(n)}(\lambda y)^n \frac{e^{-(\lambda-t)y}}{y}dy$$

$$= \frac{\lambda^n}{(\lambda-t)^n}\int_0^\infty \frac{1}{\Gamma(n)}((\lambda-t)y)^n \frac{e^{-(\lambda-t)y}}{y}dy = \frac{\lambda^n}{(\lambda-t)^n}\cdot 1$$

Gamma and Poisson Connections

- This also allows us to connect the Gamma distribution to the story of the Poisson process. In Unit 5 we saw that in a Poisson process of rate λ, the interarrival times, X_i, are i.i.d. Expo(λ) r.v.s. But the total waiting time T_n for the n^{th} arrival is the sum of the first n interarrival times.
- Therefore, $T_n \sim$ Gamma(n, λ). The interarrival times in a Poisson process are Exponential r.v.s, while the raw arrival times are Gamma r.v.s. Note: the T_n are not independent.
- A picture is worth a thousand words:



Unit 8 Outline

- Change of Variables
- Convolutions
- Gamma Distribution
- Beta Distribution
- Beta-Gamma Connections
- Order Statistics (optional)

Story of the Beta Distribution

- The Beta distribution is a generalization of the standard Unif(0,1) distribution.
- Instead of the PDF having constant density on the interval (0,1), the PDF is allowed to vary (and not necessarily symmetric on the interval).
- This will eventually be tied in with the Gamma distribution.
- Note: the Beta(1,1) is equivalent to the Unif(0,1).

Beta Distribution Definition

- A r.v. X has a ***Beta distribution***, written as $X \sim \text{Beta}(a,b)$, with parameters a and b ($a > 0$ and $b > 0$) if its PDF is:

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}$$

for $0 < x < 1$, where $\beta(a,b)$ is chosen to make the PDF integrate to 1.

- By definition, the normalizing constant, sometimes called the *Beta function*, satisfies:

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

More Beta Details

- Let $X \sim \text{Beta}(a,b)$. Find the mean and variance of X.
- Good old ***PDF recognition***! $E(X)$ first:

$$E(X) = \int_0^1 x \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}dx$$

$$= \frac{a}{(a+b)} \int_0^1 \frac{\Gamma(a+b+1)}{\Gamma(a+1)\Gamma(b)} x^{a}(1-x)^{b-1}dx = \frac{a}{(a+b)}$$

- And $E(X^2)$:

$$E(X^2) = \int_0^1 x^2 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}dx = \frac{a(a+1)}{(a+b)(a+b+1)}$$

- Thus $\text{Var}(X) = E(X^2) - [E(X)]^2 = ab/[(a+b)^2(a+b+1)]$

Unit 8 Outline

- Change of Variables
- Convolutions
- Gamma Distribution
- Beta Distribution
- Beta-Gamma Connections
- Order Statistics (optional)

Connecting Beta and Gamma

- (Bank-post office). While running errands, you need to go to the bank, and then to the post office. Let X ~ Gamma(a,λ) be your waiting time in line at the bank, and let Y ~ Gamma(b,λ) be your waiting time in line at the post office (with the same λ for both). Assume X and Y are independent. What is the joint distribution of $T = X + Y$ (your total wait at the bank and post office) and $W = X/(X+Y)$ (the fraction of your waiting time spent at the bank)?
- What should be the distribution of T?
  - T ~ Gamma($a+b$, λ)
- We'll see that the distribution of W is Beta(a,b).
- Let's derive it using our Change of Variables approach (2D)

Change of Variables multiD Example

- What's the joint PDF of X~Gamma(a,λ) and Y~Gamma(b,λ)?

$$f_{X,Y}(x,y) = \left(\frac{1}{\Gamma(a)}(\lambda x)^a e^{-\lambda x}\frac{1}{x}\right)\left(\frac{1}{\Gamma(b)}(\lambda y)^b e^{-\lambda y}\frac{1}{y}\right)$$

- Let's calculate the Jacobian matrix (Note: $t = x + y$, $w = x/(x+y)$ means that $x = tw$ and $y = t(1-w)$:

$$\frac{\partial(x,y)}{\partial(t,w)} = \begin{pmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial w} \end{pmatrix} = \begin{pmatrix} w & t \\ 1-w & -t \end{pmatrix}$$

- And its absolute determinant is $|-wt - t(1-w)| = t$.

Change of Variables multiD Example

- So let's put everything together:

$$f_{X,Y}(x,y) = f_{X,Y}(x,y)\left|\frac{\partial(x,y)}{\partial(t,w)}\right| = \left(\frac{1}{\Gamma(a)}(\lambda x)^a e^{-\lambda x}\frac{1}{x}\right)\left(\frac{1}{\Gamma(b)}(\lambda y)^b e^{-\lambda y}\frac{1}{y}\right)\cdot t$$

$$= \left(\frac{1}{\Gamma(a)}(\lambda tw)^a e^{-\lambda tw}\frac{1}{tw}\right)\left(\frac{1}{\Gamma(b)}(\lambda t(1-w))^b e^{-\lambda t(1-w)}\frac{1}{t(1-w)}\right)\cdot t$$

$$= \frac{1}{\Gamma(a)\Gamma(b)} w^{a-1}(1-w)^{b-1}(\lambda t)^{a+b} e^{-\lambda t}\frac{1}{t}$$

$$= \left(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} w^{a-1}(1-w)^{b-1}\right)\left(\frac{1}{\Gamma(a+b)}(\lambda t)^{a+b} e^{-\lambda t}\frac{1}{t}\right)$$

- For $0 < w < 1$ and $t > 0$.
- Hole cow! What distributions do we have!?

Unit 8 Outline

- Change of Variables
- Convolutions
- Gamma Distribution
- Beta Distribution
- Beta-Gamma Connections
- Order Statistics (optional)

Order Statistics

- The final transformation to consider is the one that takes n random variables $X_1,\ldots,X_n$ and sorts them in order from min to max: $\min(X_1,\ldots,X_n)\ldots\max(X_1,\ldots,X_n)$.
- These transformed r.v.s are called the *order statistics*. These are often useful when we want to worry about the distribution of extreme values.
- They often are used as summaries of an experiment: the realizations of a random experiment. Often the best or worst 2.5%, 5%, 25%, etc… are reported.

Order Statistics

- For r.v.s $X_1,\ldots,X_n$ the *order statistics* are the r.v.s , $X_{(1)},\ldots,X_{(n)}$ where:

  $X_{(1)} = \min(X_1,\ldots,X_n)$
  $X_{(2)}$ = second smallest of $X_1,\ldots,X_n$
  $\vdots$
  $X_{(n-1)}$ = second largest of $X_1,\ldots,X_n$
  $X_{(n)} = \max(X_1,\ldots,X_n)$.
- Note: $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$ by definition.
- Which order statistics is the sample median (if n is odd)?

  $X_{((n+1)/2)}$
- Are the order statistics independent or dependent?

Order Statistics

- The transformation from the original X_j to $X_{(j)}$ is not invertible. Why? Because given the resulting order statistics, say $\min(X,Y) = 1$ and $\max(X,Y) = 2$, we cannot transform back to determine what the original values were (whether X was the 1 or the 2).
- We instead need to take a direct approach.
- When dealing with the distribution of the j^{th} order statistic, it makes sense to deal with the CDF.
- Let's start with the easiest: the maximum, $X_{(n)}$, and the minimum, $X_{(1)}$.

Order Statistics

- Let $X_1,\ldots,X_n$ be i.i.d. and continuous. Let F be their CDF.
- What is the CDF of $X_{(n)}$? By definition of CDFs:

$$\begin{aligned}F_{X_{(n)}}(x) &= P(\max(X_1,\ldots,X_n) \le x)\\ &= P(X_1 \le x,\ldots,X_n \le x)\\ &= P(X_1 \le x)P(X_2 \le x)\ldots P(X_n \le x)\\ &= [F(x)]^n\end{aligned}$$

- What is the CDF of $X_{(1)}$?

$$\begin{aligned}F_{X_{(1)}}(x) &= P(\min(X_1,\ldots,X_n) \le x) = 1 - P(\min(X_1,\ldots,X_n) > x)\\ &= 1 - P(X_1 > x,\ldots,X_n > x)\\ &= 1 - [P(X_1 > x)P(X_2 > x)\ldots P(X_n > x)]\\ &= 1 - [1 - F(x)]^n\end{aligned}$$

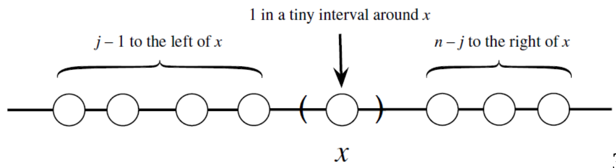
CDF of Order Statistics

- What about the event of $(X_{(j)} \le x)$?
- This means we need at least j of the X_i to fall at or below x.
- Let's define a new random variable, N, to count exactly that: the number of that fall below x.
- What distribution will N have?
- What will be the probability of success (falling at or below x) for each X_i in terms of the CDF F?
- Thus $N \sim \text{Bin}(n, p = F(x))$. So…

$$F_{X_{(j)}}(x) = P(X_{(j)} \le x) = P(\text{at least } j \text{ of the } X_i \text{ fall at or below } x)$$
$$= P(N \ge j)$$
$$= \sum_{k=j}^{n} \binom{n}{k} [F(x)]^k [1 - F(x)]^{n-k}$$

PDF of Order Statistics

- To find the PDF of the order statistic, then we just need to differentiate the CDF on the previous slide.
- But this is ugly. Instead we will take a direct approach:
- The probability that the j^{th} order statistic falls into an infinitesimal interval of length dx around x is: $f_{X_{(j)}}(x) \cdot dx$
- This can only happen if **one** of the X_i falls in this area, and **exactly *j*-1** fall below it, and **exactly *n-j*** fall above it.
- Here's the illustration:



PDF of Order Statistics

- What's the probability of this happening? Break it down:
- What the probability of exactly one of the X_i falls in this area $f(x)dx$?
  There are n choices of X_i, each with has probability $f(x)dx$
- Next, we need exactly j-1 out of the remaining n-1 to fall to the left of x. This is exactly the Binomial distribution, with probability: $\binom{n-1}{j-1} [F(x)]^{j-1} [1 - F(x)]^{n-j}$
- Put that all together, and drop the dx from both sides, and we get the PDF of $X_{(j)}$:

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

Order Statistics of Unif(0,1)

- Let i.i.d. $U_1, \ldots, U_n \sim \text{Unif}(0,1)$.
  Then for $0 \le x \le 1$, $f(x) = 1$, and $F(x) = x$.
- Then the PDF of the j^{th} order statistic, $U_{(j)}$, is (using the PDF from the previous slide):

$$f_{U_{(j)}}(x) = n \binom{n-1}{j-1} x^{j-1} (1 - x)^{n-j}$$

- What named PDF is this?
  So $U_{(j)} \sim \text{Beta}(a=j, b=n-j+1)$, with $E(U_{(j)}) = j/(n+1)$.