Proof: PMFs, PDFs, CDFs

Sophie Chowgule, Tom Coleman

Summary

Explanations as to why some PMF’s and PDF’s are valid.

*Before reading this proof sheet, it is recommended that you read* [*Guide: PMFs, PDFs, CDFs*](../studyguides/pmfspdfscdfs.qmd)*. Other recommended reading material will be said when it is needed.*

# Proof that the binomial distribution is a PMF

*Before reading this section, you may find it useful to read [Guide: The binomial theorem].*

Remember from [Guide: PMFs, PDFs, CDFs](../studyguides/pmfspdfscdfs.qmd) that the **binomial distribution** is given by the following.

|  |
| --- |
|  Binomial distribution |
| $$P\left(X=x\right)=\left(\genfrac{}{}{0pt}{}{n}{x}\right)p^{x}q^{\left(n−x\right)}=\frac{n!}{\left(n−x\right)!x!}p^{x}q^{\left(n−x\right)}$$where:* the random variable $X=x$ measures the number of success in a set of $n$ trials
	+ $x$ is number of successes
	+ $n$ is number of trials
* $p$ is the probability of success in a single trial
* $q=1−p$ is the probability of failure in a single trial
 |

Also from [Guide: PMFs, PDFs, CDFs](../studyguides/pmfspdfscdfs.qmd), the two conditions to be a valid PMF are the following:

* **Non-negativity**: The probability assigned to each possible outcome must be greater than or equal to zero, that is:

$$p\left(x\right)=P\left(X=x\right)\geq 0 for all values of x.$$

* **Honesty condition**: The sum of probabilities of all possible outcomes $x$ of a discrete random variable $X$ must be equal to one:

$$\sum\_{x}^{​}p\left(x\right)=\sum\_{x}^{​}P\left(X=x\right)=1.$$

First of all, every term in the PMF for the binomial distribution above is non-negative, and the product of non-negative numbers is non-negative, so $P\left(X=x\right)\geq 0$ for any $x$.

The honesty condition comes about because binomial distributions follow the **binomial theorem**. The binomial theorem states that:

$$\sum\_{x=0}^{n}\left(\genfrac{}{}{0pt}{}{n}{x}\right)p^{x}q^{\left(n−x\right)}=\left(p+q\right)^{n}$$

(See [Guide: The binomial theorem] for more.)

The number of successes $x$ ranges from $0$ (total failure) to $n$ (complete success). Therefore, the sum of all possible probabilities $P\left(X=x\right)$ is:

$$\sum\_{x}^{​}P\left(X=x\right)=\sum\_{x=0}^{n}\left(\genfrac{}{}{0pt}{}{n}{x}\right)p^{x}q^{\left(n−x\right)}$$

which is the left-hand side of the binomial theorem. Using the binomial theorem with $q=1−p$:

$$\sum\_{x}^{​}P\left(X=x\right)=\left(p+q\right)^{n}=\left(p+\left(1−p\right)\right)^{n}=\left(1\right)^{n}=1$$

So, the sum of the probabilities over all possible values of $x$ equals 1, satisfying the honesty condition.

# Proof that the uniform distribution is a PDF

*Before reading this section, you may find it useful to read [Guide: Introduction to integration] and [Guide: Properties of integration].*

Remember from [Guide: PMFs, PDFs, CDFs](../studyguides/pmfspdfscdfs.qmd) that the **uniform distribition** over the interval $\left[a,b\right]$ is given by the following.

|  |
| --- |
|  Uniform distribution |
| $$f\left(x\right)=\left\{\begin{matrix}\frac{1}{b−a}&if a\leq x\leq b\\0&otherwise\end{matrix}\right.$$where $a,b$ are real numbers such that $a<b$. |

Also from [Guide: PMFs, PDFs, CDFs](../studyguides/pmfspdfscdfs.qmd), the two conditions to be a valid PDF are the following:

* **Non-negativity**: The PDF $f\left(x\right)$ must be greater than or equal to zero over its entire range of possible values:

$$f\left(x\right)\geq 0 for all values of x.$$

* **Honesty condition**: The area under the entire PDF $f\left(x\right)$ must be equal to $1$, so:

$$\int\_{−\infty }^{\infty }f\left(x\right) dx=1.$$

To check if this is a valid PDF, you need to confirm that it satisfies these two key conditions.

**Non-negativity**: $f\left(x\right)\geq 0$ for all values of $x$, as $f\left(x\right)=\frac{1}{b−a}$ in $\left[a,b\right]$ and $0$ otherwise.

**Honesty**: To satisfy the honesty condition, the integral of the PDF over the interval $\left[a,b\right]$ must equal $1$. Using the properties of integration, you can split the integral into three parts along the lines of the PDF:

$$\int\_{−\infty }^{\infty }f\left(x\right) dx=\int\_{−\infty }^{a}f\left(x\right) dx+\int\_{a}^{b}f\left(x\right) dx+\int\_{b}^{\infty }f\left(x\right) dx$$

Using the definition of $f\left(x\right)$ on these intervals gives

$$\int\_{−\infty }^{a}f\left(x\right) dx+\int\_{a}^{b}f\left(x\right) dx+\int\_{b}^{\infty }f\left(x\right) dx=\int\_{−\infty }^{a}0 dx+\int\_{a}^{b}\frac{1}{b−a} dx+\int\_{b}^{\infty }0 dx$$

Since the integral of $0$ over any limits is zero, this reduces to

$$\int\_{−\infty }^{\infty }f\left(x\right) dx=0+\int\_{a}^{b}\frac{1}{b−a} dx+0=\int\_{a}^{b}\frac{1}{b−a} dx$$

Working out this integral dives

$$\int\_{a}^{b}\frac{1}{b−a} dx=\frac{1}{b−a}\int\_{a}^{b}1 dx=\frac{1}{b−a}[x]\_{a}^{b}=\frac{1}{b−a}\left(b−a\right)=1$$

And so you can see that all uniform distributions are valid PDFs.

# Proof that the normal distribution is a PDF

*Before reading this section, you may find it useful to read [Guide: Properties of integration], [Guide: Integration by substitution], [Guide: Introduction to double integration], and [Guide: Co-ordinate changes in double integration].*

Remember from [Guide: PMFs, PDFs, CDFs](../studyguides/pmfspdfscdfs.qmd) that the **normal distribution** is given by the following.

|  |
| --- |
|  Normal distribution |
| $$f\left(x\right)=\frac{1}{σ\sqrt{2π}}exp\left(−\frac{1}{2}\left(\frac{x−μ}{σ}\right)^{2}\right)$$where $μ,σ$ are real numbers such that $σ>0$. (Here, $μ$ is the mean and $σ$ is the standard deviation.) |

To check if this is a valid PDF, you need to confirm that it satisfies the two key conditions.

**Non-negativity**: As an exponential function, $exp\left(−\frac{1}{2}\left(\frac{x−μ}{σ}\right)^{2}\right)>0$, and $1/σ\sqrt{2π}>0$ as $σ>0$. So $f\left(x\right)>0$.

**Honesty**: Here’s the fun part.

The idea is to show that this integral $I$, given by

$$I=\int\_{−\infty }^{\infty }f\left(x\right) dx=\int\_{−\infty }^{\infty }\frac{1}{σ\sqrt{2π}}exp\left(−\frac{1}{2}\left(\frac{x−μ}{σ}\right)^{2}\right) dx$$

is equal to $1$. To tackle this integral, it needs to look a little nicer; you can use integration by substitution to do this (see [Guide: Integration by substitution]). Let $u=\frac{x−μ}{σ\sqrt{2}}$. Then $\frac{du}{dx}=\frac{1}{σ\sqrt{2}}$, and so $dx=σ\sqrt{2} du$. As $x\rightarrow \pm \infty $, it follows that $u\rightarrow \pm \infty $. Since $u^{2}=\frac{1}{2}\left(\frac{x−μ}{σ}\right)^{2}$, the integral becomes

$$\int\_{−\infty }^{\infty }\frac{1}{σ\sqrt{2π}}exp\left(−\frac{1}{2}\left(\frac{x−μ}{σ}\right)^{2}\right) dx=\int\_{−\infty }^{\infty }\frac{σ\sqrt{2}}{σ\sqrt{2π}}e^{−u^{2}} du=\frac{1}{\sqrt{π}}\int\_{−\infty }^{\infty }e^{−u^{2}} du$$

Next, you can use the fact that $exp\left(−u^{2}\right)$ is an even function to change the limits. Using the property of even function about symmetric limits (see [Guide: Properties of integration]), the integral becomes

$$\frac{1}{\sqrt{π}}\int\_{−\infty }^{\infty }e^{−u^{2}} du=\frac{2}{\sqrt{π}}\int\_{0}^{\infty }e^{−u^{2}} du=I$$

All that you have done so far has not changed the value of the integral, so this is still equal to $I$. Now, the choice of variables in an integral doesn’t matter, so $I=\frac{2}{\sqrt{π}}\int\_{0}^{\infty }e^{−v^{2}} dv$ as well. Multiplying both together gives

$$I^{2}=\frac{4}{π}\left(\int\_{0}^{\infty }e^{−u^{2}} du\right)\left(\int\_{0}^{\infty }e^{−v^{2}} dv\right)$$

Now, the variables here are independent, so you can combine this into a double integral. Doing this gives

$$I^{2}=\frac{1}{π}\left(\int\_{0}^{\infty }e^{−u^{2}} du\right)\left(\int\_{0}^{\infty }e^{−v^{2}} dv\right)=\frac{1}{π}\int\_{0}^{\infty }\int\_{0}^{\infty }e^{−\left(u^{2}+v^{2}\right)} du dv$$

You can now change the co-ordinates to polar co-ordinates (see [Guide: Changing co-ordinates in double integrals] for more). By setting $u=rcos\left(θ\right)$ and $v=rsin\left(θ\right)$, it follows that $u^{2}+v^{2}=r^{2}$. The region of integration is $0\leq u<\infty $ and $0\leq v<\infty $, which corresponds to the first quadrant of the plane; this is represented in polar co-ordinates by $0\leq r<\infty $ and $0\leq θ\leq π/2$. Finally, $du dv$ becomes $rdr dθ$ by using the Jacobian. Therefore, the integral becomes

$$I^{2}=\frac{4}{π}\int\_{0}^{\infty }\int\_{0}^{\infty }e^{−\left(u^{2}+v^{2}\right)} du dv=\frac{4}{π}\int\_{0}^{π/2}\int\_{0}^{\infty }re^{−r^{2}} dr dθ$$

Now you can evaluate this double integral. The derivative of $e^{−r^{2}}$ with respect to $r$ is $−2re^{−r^{2}}$; so that means that the integral of $re^{−r^{2}}$ is $−\frac{1}{2}e^{−r^{2}}$ (you can get this result by substitution if you wanted). Using the fact that $e^{−r^{2}}$ is equal to $1$ when $r=0$ and tends to $0$ as $r$ tends to infinity, you can get

$$I^{2}=\frac{4}{π}\int\_{0}^{π/2}\int\_{0}^{\infty }re^{−r^{2}} dr dθ=\frac{4}{π}\int\_{0}^{π/2}\left[−\frac{1}{2}e^{−r^{2}}\right]\_{0}^{\infty } dθ=\frac{4}{π}\int\_{0}^{π/2}\frac{1}{2} dθ$$

Evaluating this final integral gives

$$I^{2}=\frac{4}{π}\int\_{0}^{π/2}\frac{1}{2} dθ=\frac{4}{π}\left[\frac{θ}{2}\right]\_{0}^{π/2}=\frac{4}{π}⋅\frac{π}{4}=1$$

So $I^{2}=1$, implying that $I=\pm 1$. But $I$ cannot be $−1$, as $f\left(x\right)$ is a positive function and the integral of a positive function is always positive. So $I=1$ and therefore the normal distribution really is a PDF.

# Further reading

[Guide: PMFs, PDFs, CDFs](../studyguides/pmfspdfscdfs.qmd)

[Questions: PMFs, PDFs, CDFs](../questions/qs-pmfspdfscdfs.qmd)

## Version history and licensing

v1.0: initial version created in 04/25 by tdhc and Sophie Chowgule as part of a University of St Andrews VIP project.

[This work is licensed under CC BY-NC-SA 4.0.](https://creativecommons.org/licenses/by-nc-sa/4.0/?ref=chooser-v1)