

PROBABILITY DISTRIBUTIONS: MEANS AND VARIANCES

RADON ROSBOROUGH

<https://intuitiveexplanations.com/math/probability-distributions-means-and-variances/>

This document is a collection of derivations for the formulas for the means and variances of seven different probability distributions commonly encountered in statistics.

1. MEAN OF A LINEAR FUNCTION OF A DISCRETE PROBABILITY DISTRIBUTION

$$\begin{aligned}
\mu_{aX+b} &= \sum_{x=0}^{\infty} (ax + b)p_X(x) \\
&= a \sum_{x=0}^{\infty} xp_X(x) + b \sum_{x=0}^{\infty} p_X(x) \\
&= a\mu_X + b.
\end{aligned}$$

2. VARIANCE OF A LINEAR FUNCTION OF A DISCRETE PROBABILITY DISTRIBUTION

$$\begin{aligned}
\sigma_{aX+b}^2 &= \sum_{x=0}^{\infty} (ax + b - \mu_{aX+b})^2 p_X(x) \\
&= \sum_{x=0}^{\infty} (ax + b - a\mu_X - b)^2 p_X(x) \\
&= a^2 \sum_{x=0}^{\infty} (x - \mu_X)^2 p_X(x) \\
&= a^2 \sigma_X^2.
\end{aligned}$$

3. MEAN OF A SUM OF DISCRETE PROBABILITY DISTRIBUTIONS

$$\begin{aligned}
\mu_{X+Y} &= \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} (x + y)p_X(x)p_Y(y) \\
&= \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} xp_X(x)p_Y(y) + \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} yp_X(x)p_Y(y) \\
&= \sum_{x=0}^{\infty} \left[xp_X(x) \sum_{y=0}^{\infty} p_Y(y) \right] + \sum_{x=0}^{\infty} \left[p_X(x) \sum_{y=0}^{\infty} yp_Y(y) \right] \\
&= \sum_{x=0}^{\infty} xp_X(x) + \mu_Y \sum_{x=0}^{\infty} p_X(x) \\
&= \mu_X + \mu_Y.
\end{aligned}$$

4. VARIANCE OF A SUM OF DISCRETE PROBABILITY DISTRIBUTIONS

$$\begin{aligned}
\sigma_{x+y}^2 &= \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} (x + y - \mu_{X+Y})^2 p_X(x)p_Y(y) \\
&= \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} (x + y - \mu_X - \mu_Y)^2 p_X(x)p_Y(y)
\end{aligned}$$

$$\begin{aligned}
 &= \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} (x - \mu_X)^2 p_X(x) p_Y(y) + \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} (y - \mu_Y)^2 p_X(x) p_Y(y) \\
 &\quad + 2 \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} (x - \mu_X)(y - \mu_Y) p_X(x) p_Y(y) \\
 &= \sum_{x=0}^{\infty} \left[(x - \mu_X)^2 p_X(x) \sum_{y=0}^{\infty} p_Y(y) \right] + \sum_{x=0}^{\infty} \left[p_X(x) \sum_{y=0}^{\infty} (y - \mu_Y)^2 p_Y(y) \right] \\
 &\quad + 2 \sum_{x=0}^{\infty} \left\{ (x - \mu_X) p_X(x) \left[\sum_{y=0}^{\infty} y p_Y(y) - \mu_Y \sum_{y=0}^{\infty} p_Y(y) \right] \right\} \\
 &= \sum_{x=0}^{\infty} (x - \mu_X)^2 p_X(x) + \sigma_Y^2 \sum_{x=0}^{\infty} p_X(x) + 2 \sum_{x=0}^{\infty} (x - \mu_X) p_X(x) (\mu_Y - \mu_Y) \\
 &= \sigma_X^2 + \sigma_Y^2.
 \end{aligned}$$

5. MEAN OF A LINEAR FUNCTION OF A CONTINUOUS PROBABILITY DISTRIBUTION

$$\begin{aligned}
 \mu_{aX+b} &= \int_{-\infty}^{\infty} (ax + b) p_X(x) dx \\
 &= a \int_{-\infty}^{\infty} x p_X(x) dx + b \int_{-\infty}^{\infty} p_X(x) dx \\
 &= a\mu_X + b.
 \end{aligned}$$

6. VARIANCE OF A LINEAR FUNCTION OF A CONTINUOUS PROBABILITY DISTRIBUTION

$$\begin{aligned}
 \sigma_{aX+b}^2 &= \int_{-\infty}^{\infty} (ax + b - \mu_{aX+b})^2 p_X(x) dx \\
 &= \int_{-\infty}^{\infty} (ax + b - a\mu_X - b)^2 p_X(x) dx \\
 &= a^2 \int_{-\infty}^{\infty} (x - \mu_X)^2 p_X(x) dx \\
 &= a^2 \sigma_X^2.
 \end{aligned}$$

7. MEAN OF A SUM OF CONTINUOUS PROBABILITY DISTRIBUTIONS

$$\begin{aligned}
 \mu_{X+Y} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y) p_X(x) p_Y(y) dy dx \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x p_X(x) p_Y(y) dy dx + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y p_X(x) p_Y(y) dy dx \\
 &= \int_{-\infty}^{\infty} \left[x p_X(x) \int_{-\infty}^{\infty} p_Y(y) dy \right] dx + \int_{-\infty}^{\infty} \left[p_X(x) \int_{-\infty}^{\infty} y p_Y(y) dy \right] dx \\
 &= \int_{-\infty}^{\infty} x p_X(x) dx + \mu_Y \int_{-\infty}^{\infty} p_X(x) dx
 \end{aligned}$$

$$= \mu_X + \mu_Y.$$

8. VARIANCE OF A SUM OF CONTINUOUS PROBABILITY DISTRIBUTIONS

$$\begin{aligned}
\sigma_{X+Y}^2 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+y-\mu_{X+Y})^2 p_X(x)p_Y(y) dy dx \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+y-\mu_X-\mu_Y)^2 p_X(x)p_Y(y) dy dx \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x-\mu_X)^2 p_X(x)p_Y(y) dy dx + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y-\mu_Y)^2 p_X(x)p_Y(y) dy dx \\
&\quad + 2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y)p_X(x)p_Y(y) dy dx \\
&= \int_{-\infty}^{\infty} \left[(x-\mu_X)^2 p_X(x) \int_{-\infty}^{\infty} p_Y(y) dy \right] dx + \int_{-\infty}^{\infty} \left[p_X(x) \int_{-\infty}^{\infty} (y-\mu_Y)^2 p_Y(y) dy \right] dx \\
&\quad + 2 \int_{-\infty}^{\infty} \left\{ (x-\mu_X)p_X(x) \left[\int_{-\infty}^{\infty} yp_Y(y) dy - \mu_Y \int_{-\infty}^{\infty} p_Y(y) dy \right] \right\} dx \\
&= \int_{-\infty}^{\infty} (x-\mu_X)^2 p_X(x) dx + \sigma_Y^2 \int_{-\infty}^{\infty} p_X(x) dx + 2 \int_{-\infty}^{\infty} (x-\mu_X)p_X(x)(\mu_Y - \mu_Y) dx \\
&= \sigma_X^2 + \sigma_Y^2.
\end{aligned}$$

9. MEAN OF A NORMAL DISTRIBUTION

$$\begin{aligned}
\mu_{\text{Normal}[\mu, \sigma]} &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x \exp\left(-\frac{1}{2} \left[\frac{x-\mu}{\sigma}\right]^2\right) dx \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (u+\mu) \exp\left(-\frac{1}{2} \left[\frac{u}{\sigma}\right]^2\right) du, & u = x - \mu \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} u \exp\left(-\frac{1}{2} \left[\frac{u}{\sigma}\right]^2\right) du + \frac{\mu}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2} \left[\frac{u}{\sigma}\right]^2\right) du \\
&= \frac{\mu}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-v^2} dv, & v = \frac{u}{\sqrt{2}\sigma} \\
&= \mu.
\end{aligned}$$

10. VARIANCE OF A NORMAL DISTRIBUTION

$$\begin{aligned}
\sigma_{\text{Normal}[\mu, \sigma]}^2 &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-\mu_{\text{Normal}[\mu, \sigma]})^2 \exp\left(-\frac{1}{2} \left[\frac{x-\mu}{\sigma}\right]^2\right) dx \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x-\mu)^2 \exp\left(-\frac{1}{2} \left[\frac{x-\mu}{\sigma}\right]^2\right) dx \\
&= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} y^2 e^{-y^2} dy, & y = \frac{x-\mu}{\sqrt{2}\sigma}
\end{aligned}$$

$$\begin{aligned}
&= \frac{2\sigma^2}{\sqrt{\pi}} \left\{ \left[-\frac{1}{2}ye^{-y^2} \right]_{-\infty}^{\infty} + \frac{1}{2} \int_{-\infty}^{\infty} e^{-y^2} dy \right\}, & u = y, dv = ye^{-y^2} \\
&= \frac{\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-y^2} dy \\
&= \sigma^2.
\end{aligned}$$

11. MEAN OF A BINOMIAL DISTRIBUTION

$$\begin{aligned}
\mu_{\text{Binomial}[n,p]} &= \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^n k \left[\frac{n!}{k!(n-k)!} \right] p^k (1-p)^{n-k} \\
&= \sum_{k=1}^n n \left[\frac{(n-1)!}{(k-1)!(n-k)!} \right] p^k (1-p)^{n-k} \\
&= n \sum_{k=0}^{n-1} \left[\frac{(n-1)!}{k!(n-1-k)!} \right] p^{k+1} (1-p)^{n-1-k} \\
&= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} \\
&= np(p+1-p)^{n-1} \\
&= np.
\end{aligned}$$

12. VARIANCE OF A BINOMIAL DISTRIBUTION

$$\begin{aligned}
\sigma_{\text{Binomial}[n,p]}^2 &= \sum_{k=0}^n (k - \mu_{\text{Binomial}[n,p]})^2 \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=0}^n (k - np)^2 \binom{n}{k} p^k (1-p)^{n-k} \\
&= n^2 p^2 \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} - 2np \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} + \sum_{k=0}^n k^2 \binom{n}{k} p^k (1-p)^{n-k} \\
&= n^2 p^2 (p+1-p)^n - 2np \mu_{\text{Binomial}[n,p]} + \sum_{k=1}^n k^2 \binom{n}{k} p^k (1-p)^{n-k} \\
&= n^2 p^2 - 2n^2 p^2 + \sum_{k=1}^n k \left[\frac{n!}{(k-1)!(n-k)!} \right] p^k (1-p)^{n-k} \\
&= -n^2 p^2 + n \sum_{k=1}^n k \left[\frac{(n-1)!}{(k-1)!(n-k)!} \right] p^k (1-p)^{n-k}
\end{aligned}$$

$$\begin{aligned}
&= -n^2 p^2 + n \sum_{k=0}^{n-1} (k+1) \left[\frac{(n-1)!}{k!(n-1-k)!} \right] p^{k+1} (1-p)^{n-1-k} \\
&= -n^2 p^2 + np \sum_{k=0}^{n-1} (k+1) \binom{n-1}{k} p^k (1-p)^{n-1-k} \\
&= -n^2 p^2 + np \left[\sum_{k=0}^{n-1} k \binom{n-1}{k} p^k (1-p)^{n-1-k} + \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} \right] \\
&= -n^2 p^2 + np [\mu_{\text{Binomial}[n-1,p]} + (p+1-p)^{n-1}] \\
&= -n^2 p^2 + n(n-1)p^2 + np \\
&= np - np^2 \\
&= np(1-p).
\end{aligned}$$

13. MEAN OF A GEOMETRIC DISTRIBUTION

$$\begin{aligned}
\mu_{\text{Geometric}[p]} &= \sum_{n=1}^{\infty} np(1-p)^{n-1} \\
&= p \sum_{n=1}^{\infty} n(1-p)^{n-1} \\
\sum_{n=0}^{\infty} x^n &= \frac{1}{1-x} \\
\sum_{n=1}^{\infty} nx^{n-1} &= \frac{1}{(1-x)^2} \\
\sum_{n=1}^{\infty} n(1-p)^{n-1} &= \frac{1}{p^2} \\
\mu_{\text{Geometric}[p]} &= \frac{1}{p}.
\end{aligned}$$

14. VARIANCE OF A GEOMETRIC DISTRIBUTION

$$\begin{aligned}
\sigma_{\text{Geometric}[p]}^2 &= \sum_{n=1}^{\infty} (n - \mu_{\text{Geometric}[p]})^2 p(1-p)^{n-1} \\
&= \sum_{n=1}^{\infty} \left(n - \frac{1}{p} \right)^2 p(1-p)^{n-1} \\
&= \sum_{n=1}^{\infty} \left(\frac{1}{p^2} - \frac{2n}{p} + n^2 \right) p(1-p)^{n-1} \\
&= \frac{1}{p} \sum_{n=1}^{\infty} (1-p)^{n-1} - 2 \sum_{n=1}^{\infty} n(1-p)^{n-1} + p \sum_{n=1}^{\infty} n^2 (1-p)^{n-1}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{p} \left[\frac{1}{1 - (1 - p)} \right] - \frac{2}{p} \sum_{n=1}^{\infty} np(1 - p)^{n-1} + p \sum_{n=1}^{\infty} n^2(1 - p)^{n-1} \\
&= \frac{1}{p^2} - \frac{2\mu_{\text{Geometric}[p]}}{p} + p \sum_{n=1}^{\infty} n^2(1 - p)^{n-1} \\
&= -\frac{1}{p^2} + p \sum_{n=1}^{\infty} n^2(1 - p)^{n-1} \\
\sum_{n=0}^{\infty} x^n &= \frac{1}{1 - x} \\
\sum_{n=1}^{\infty} nx^{n-1} &= \frac{1}{(1 - x)^2} \\
\sum_{n=1}^{\infty} nx^n &= \frac{x}{(1 - x)^2} \\
\sum_{n=1}^{\infty} n^2x^n &= \frac{(1 - x)^2 + 2x(1 - x)}{(1 - x)^4} \\
&= \frac{1 - x + 2x}{(1 - x)^3} \\
&= \frac{1 + x}{(1 - x)^3} \\
\sum_{n=1}^{\infty} n^2(1 - p)^{n-1} &= \frac{2 - p}{p^3} \\
\sigma_{\text{Geometric}[p]}^2 &= -\frac{1}{p^2} + \frac{2 - p}{p^2} \\
&= \frac{1 - p}{p^2}.
\end{aligned}$$