Week 10

Nov. 5 – Nov. 9

Lecture 26. Sum of independent Normal random variables
The most important density function

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

There is a special notation for the density $\phi(x)$ and for the cumulative distribution function $\Phi(x)$.

It is even not easy to see why

$$\int \phi(x)\,dx = 1?$$

Trick:

$$\begin{aligned}\left(\int \phi(x)\,dx\right)^2 &= \int \phi(x)\,dx \int \phi(y)\,dy \\ &= \frac{1}{2\pi}\int\int \exp\left(-\frac{1}{2}\left(x^2+y^2\right)\right) dxdy\end{aligned}$$

Now switch to polar coordinates

$$x = r\cos\theta,\ y = r\sin\theta,$$

$$\left(\int \phi(x)\,dx\right)^2 = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{1} \exp\left(-\frac{1}{2}r^2\right) rdrd\theta = \frac{1}{2\pi}\int_0^{2\pi}\left[-\exp\left(-\frac{1}{2}r^2\right)\right]_0^{\infty} d\theta = 1.$$

Question:

$$EX?$$

Question:

$$\begin{aligned}EX^2 &= ? \\ Var(X) &= ?\end{aligned}$$

Question: Let $Y = \mu + X\sigma$ with $X \sim N(0,1)$. Then

$$\begin{aligned}EY &= ? \\ Var(Y) &= ? \\ f_Y(x) &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\end{aligned}$$

More generally, we say $Y \sim N\left(\mu,\sigma^2\right)$ if $\frac{Y-\mu}{\sigma} \sim N(0,1)$.

Normal distribution – An approximation of Binomial by Abraham de Moivre

Let $X \sim Binomial\left(n,p\right)$. It can be shown that

$$P\left(a \leq \frac{X-np}{\sqrt{npq}} \leq b\right) \approx \int_a^b \phi\left(x\right)dx$$

Sum of independent Normal random variables

Let $X \sim N\left(0,\sigma_X^2\right)$ and $Y \sim N\left(0,\sigma_Y^2\right)$ be independent. The joint density of $\left(X,Y\right)$ is

$$f_{X,Y}\left(x,y\right) = \frac{1}{2\pi\sigma_X\sigma_Y}\exp\left(-\frac{x^2}{2\sigma_X^2} - \frac{y^2}{2\sigma_Y^2}\right).$$

What is joint density of $\left(Y,Z\right)$ with $Z = X+Y$?

$$\begin{aligned} f_{X,Z}\left(x,z\right) &= f_{Z|Y}\left(z|y\right)f_Y\left(y\right) \\ &= \frac{1}{2\pi\sigma_X\sigma_Y}\exp\left(-\frac{y^2}{2\sigma_Y^2} - \frac{\left(z-y\right)^2}{2\sigma_X^2}\right) \end{aligned}$$

Theorem. More generally, let f_X and f_Y be the densities of two independent random variables X and Y respectively. Let $Z = X+Y$, then

$$\begin{aligned} f_{X,Z}\left(x,z\right) &= f_{Z|X}\left(z|y\right)f_Y\left(y\right) \\ &= f_X\left(z-y\right)f_Y\left(y\right) \end{aligned}$$

and the marginal density of Z is

$$f_Z\left(z\right) = \int f_X\left(z-y\right)f_Y\left(y\right)dy$$

Definition. The convolution $f*g$ of f and g is the function given by

$$f*g = \int f\left(z-y\right)g\left(y\right)dy.$$

Let $X \sim N\left(0,\sigma_X^2\right)$ and $Y \sim N\left(0,\sigma_Y^2\right)$ be independent. The marginal density of Z is

$$\begin{aligned} f_Z\left(z\right) &= \frac{1}{2\pi\sigma_X\sigma_Y}\int \exp\left(-\frac{y^2}{2\sigma_Y^2} - \frac{\left(z-y\right)^2}{2\sigma_X^2}\right)dy \\ &= \frac{1}{2\pi\sigma_X\sigma_Y}\int \exp\left(-\frac{y^2}{2}\left(\frac{1}{\sigma_Y^2}+\frac{1}{\sigma_X^2}\right) + \frac{zy}{\sigma_X^2} - \frac{z^2}{2\sigma_X^2}\right)dy \end{aligned}$$

Note that

$$\int \exp\left[-a\frac{\left(y-b\right)^2}{2}+c\right]dy = \frac{\sqrt{2\pi}}{\sqrt{a}}e^c$$

i.e.,

$$\int \exp\left[-\frac{a}{2}y^2 + yab - \frac{ab^2}{2} + c\right]dy = \frac{2\pi}{\sqrt{a}}e^c.$$

Let

$$\begin{aligned}
a &= \frac{1}{\sigma_Y^2} + \frac{1}{\sigma_X^2} = \frac{\sigma_X^2 + \sigma_Y^2}{\sigma_X^2 \sigma_Y^2} \\
b &= \frac{z\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2} \\
c &= -\frac{z^2}{2\sigma_X^2} + \frac{1}{2}\frac{\sigma_X^2 + \sigma_Y^2}{\sigma_X^2 \sigma_Y^2} \cdot \left(\frac{z\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2}\right)^2 \\
&= -\frac{z^2}{2\sigma_X^2} + \frac{1}{2\sigma_X^2} \cdot \frac{z^2\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2} = -\frac{1}{2}\frac{z^2}{\sigma_X^2 + \sigma_Y^2}
\end{aligned}$$

then

$$\begin{aligned}
f_Z(z) &= \frac{\sigma_X \sigma_Y}{\sqrt{2\pi}\sigma_X \sigma_Y \cdot \sqrt{\sigma_X^2 + \sigma_Y^2}} \exp\left(-\frac{1}{2}\frac{z^2}{\sigma_X^2 + \sigma_Y^2}\right) \\
&= \frac{\sigma_X \sigma_Y}{\sqrt{2\pi}\sqrt{\sigma_X^2 + \sigma_Y^2}} \exp\left(-\frac{1}{2}\frac{z^2}{\sigma_X^2 + \sigma_Y^2}\right).
\end{aligned}$$

Thus

$$X \sim N\left(0, \sigma_X^2\right), Y \sim N\left(0, \sigma_Y^2\right) \text{ independent} \implies X + Y \sim N\left(0, \sigma_X^2 + \sigma_Y^2\right)$$

Lecture 27. Exponential and Gamma density – Poisson waiting time

Review: Distribution of $\mathbf{X+Y}$

Let X and Y be two independent random variables, and f_X and f_Y be the densities of X and Y respectively. Let $Z = X + Y$. What is distribution of Z? For example,

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},\ f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}.$$

The joint density of (X, Y) is

$$f_{X,Y}(x, y) = f_X(x) f_Y(y).$$

Then

$$\begin{aligned} P(X+Y \le t) &= P(-\infty < Y < \infty, X+Y \le t) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{t-y} f_X(x) f_Y(y)\, dx dy \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{t-y} f_X(x)\, dx \right] f_Y(y)\, dy. \end{aligned}$$

Let $G(t) = \int_{-\infty}^{t-y} f_X(x)\, dx$. We write

$$F_{X+Y}(t) = P(X+Y \le t) = \int_{-\infty}^{\infty} G(t) f_Y(y)\, dy.$$

This implies

$$F'_{X+Y}(t) = \int_{-\infty}^{\infty} G'(t) f_Y(y)\, dy$$

i.e.,

$$f_{X+Y}(t) = \int_{-\infty}^{\infty} f_X(t-y) f_Y(y)\, dy$$

When $X \sim N(0,1)$ and $Y \sim N(0,1)$, then

$$f_{X+Y}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{(t-y)^2}{2} - \frac{y^2}{2} \right) dy = f_{X+Y}(t) = \frac{1}{2\pi} e^{-t^2/4} \int_{-\infty}^{\infty} \exp\left(-(y-t/2)^2 \right) dy = ?.$$

Definition. The convolution $f * g$ of f and g is the function given by

$$f * g = \int f(z-y) g(y)\, dy.$$

By this definition

$$f_{X+Y} = f_X * f_Y(y).$$

Exponential distribution

$$f(x) = \lambda \exp(-\lambda x),\ x \ge 0.$$

Connection to Poisson distribution.

We consider the number of typos during a class. In a time interval of length t, assume that the number of typos Y has a Poisson distribution with expectation λt, i.e.,

$$Y \sim Poisson\left(\lambda t\right).$$

Let X_1 be the waiting time for the first typo. What is the distribution of X_1? Answer

$$P\left(X_1 > t\right) = 1 - \exp\left(-\lambda t\right)$$

and the density is

$$f_{X_1}\left(t\right) = \lambda \exp\left(-\lambda t\right), \, t > 0.$$

Let X be the waiting time for the $n-th$ typo, then

$$P\left(X > t\right) = 1 - \exp\left(-\lambda t\right)\left(1 + \frac{\lambda t}{1!} + \ldots + \frac{\left(\lambda t\right)^{n-1}}{(n-1)!}\right).$$

and the density is

$$\begin{aligned}
& \lambda \exp\left(-\lambda t\right)\left(1 + \frac{\lambda t}{1!} + \ldots + \frac{\left(\lambda t\right)^{n-1}}{(n-1)!}\right) - \exp\left(-\lambda t\right)\left(\lambda + \ldots + \frac{\lambda\left(\lambda t\right)^{n-2}}{(n-2)!}\right) \\
= & \lambda \exp\left(-\lambda t\right)\left(1 + \frac{\lambda t}{1!} + \ldots + \frac{\left(\lambda t\right)^{n-1}}{(n-1)!}\right) - \lambda \exp\left(-\lambda t\right)\left(1 + \ldots + \frac{\left(\lambda t\right)^{n-2}}{(n-2)!}\right) \\
= & \lambda \exp\left(-\lambda t\right)\frac{\left(\lambda t\right)^{n-1}}{(n-1)!} = \frac{\lambda^n}{\Gamma\left(n\right)} t^{n-1} \exp\left(-\lambda t\right)
\end{aligned}$$

where $\Gamma\left(n\right) = (n-1)!$.

Definition. $X_1 \sim$Exponential(λ),

$$f_{X_1}\left(t\right) = \lambda \exp\left(-\lambda t\right), \, t > 0.$$

$X \sim$Gamma(α, λ)

$$f_X\left(t\right) = \frac{\lambda^\alpha}{\Gamma\left(\alpha\right)} t^{\alpha-1} \exp\left(-\lambda t\right), \, t > 0.$$

Theorem.

$$X \sim Gamma\left(\alpha_1, \lambda\right), \, Y \sim Gamma\left(\alpha_2, \lambda\right) \Rightarrow X + Y \sim Gamma\left(\alpha_1 + \alpha_2, \lambda\right)$$

Proof. Applying $f_{X+Y} = f_X * f_Y\left(y\right)$.

Question:

$$\int_0^\infty t^{n-1} \exp\left(-\lambda t\right) = \frac{\Gamma\left(n\right)}{\lambda^n}?$$

Question:

$$\begin{aligned} EX_1 &= \int_0^\infty t\lambda \exp(-\lambda t)\, dt = \lambda \int_0^\infty t \exp(-\lambda t)\, dt =? \\ EX_1^2 &= \int_0^\infty t^2 \lambda \exp(-\lambda t)\, dt = \lambda \int_0^\infty t^2 \exp(-\lambda t)\, dt =? \end{aligned}$$

Let X be exponentially distributed with expectation μ and variance σ^2. Then

$$\begin{aligned} \mu &= 1/\lambda? \\ \sigma^2 &= 1/\lambda^2? \end{aligned}$$

Question:

$$\begin{aligned} EX &= ? \\ Var(X) &= ? \end{aligned}$$

Lecture 28. Transformation of random variables.
A quick summary table

X_i	$\sum_{i=1}^{n} X_i$	Remark
Bernoulli(p)	Binomial(n,p)	Y_1 ∼Binomial(n_1,p), Y_2 ∼Binomial(n_2,p), Y_1+Y_2 ∼?
Geometric(p)	Negative Binomial(n,p)	Y_1 ∼NB(n_1,p), Y_2 ∼NB(n_2,p), Y_1+Y_2 ∼?
Poisson(λ)	Poisson($n\lambda$)	Y_1 ∼Poisson(λ_1), Y_2 ∼Poisson(λ_2), Y_1+Y_2 ∼?
Exponential(λ)	Gamma(n,λ)	X ∼Gamma(α_1,λ) , Y ∼Gamma(α_2,λ), Y_1+Y_2 ∼?
Normal$\left(\mu,\sigma^2\right)$	Normal$\left(n\mu,n\sigma^2\right)$	Y_1 ∼Normal$\left(\mu_1,\sigma_1^2\right)$, Y_2 ∼Normal$\left(\mu_2,\sigma_2^2\right)$, Y_1+Y_2 ∼?

Transformation.

Let X be a continuous random variable with density $f_X(x)$. Let $h(x)$ be a strictly increasing function on the range of X. Define $Y=h(X)$. Then cdf is

$$F_Y(y)=F_X\left(h^{-1}(y)\right)$$

and pdf is

$$f_Y(y)=f_X\left(h^{-1}(y)\right)\cdot\left(h^{-1}(y)\right)'$$

Similarly, if Let $h(x)$ be a strictly decreasing function on the range of X. Define $Y=h(X)$. Then cdf is

$$F_Y(y)=1-F_X\left(h^{-1}(y)\right)$$

and pdf is

$$f_Y(y)=-f_X\left(h^{-1}(y)\right)\cdot\left(h^{-1}(y)\right)'$$

Question: Let X be a continuous random variable with cdf $F_X(x)$. Define $Y=F_X(X)$. what are cdf and pdf of Y?

All connected to uniform.

Let U denote a uniform random variable on $[0,1]$, i.e.

$$f_U(x)=1,\ 0<x<1.$$

Let X be a continuous random variable with cdf $F_X(x)$. We can built the connection of X and Z by defining

$$Y=F_X^{-1}(U).$$

Then X and Y are identically distributed!

Example (not monotone): $X\sim N(0,1)$. Define $Y=X^2$. Find cdf and pdf of Y.

$$Y\sim \text{Gamma}(1/2,1/2)?$$

Note that

$$\begin{aligned}F_{X^2}(t) &= P\left(X^2\le t\right)=P\left(-\sqrt{t}\le X\le\sqrt{t}\right)\\ &= 2\Phi\left(\sqrt{t}\right)-1\end{aligned}$$

then

$$f_{Z_1^2}(t)=\frac{1}{\sqrt{2\pi}}t^{1/2-1}\exp(-t/2).$$

Lecture 29. Chi-square, t and Cauchy
Distribution of X/Y
Let $Z_1, Z_2, \ldots, Z_n$ be i.i.d. $N(0,1)$.
Chi-square.

$$Y = Z_1^2 + \ldots + Z_n^2 \sim \chi_n^2$$

What is pdf of $Y = Z_1^2 + \ldots + Z_n^2$? Recall Z_1^2 ~Gamma$(1/2, 1/2)$. Then

$$Y \sim Gamma\,(n/2, 1/2)$$

and

$$f_Y(y) = \frac{1}{\Gamma(n/2)\, 2^n} t^{n/2-1} \exp(-t/2).$$

Cauchy.
What is the distribution of $Z_1/|Z_2|$?

$$f_{Z_1/|Z_2|}(t) = \frac{1}{\pi}\frac{1}{1+x^2}$$

Let X and Y $(Y > 0)$ be two independent random variables, and f_X and f_Y be the densities of X and Y respectively. Let $Z = X/Y$. What is distribution of Z? For example,

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},\ f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}.$$

The joint density of (X, Y) is

$$f_{X,Y}(x,y) = f_X(x)\, f_Y(y).$$

Then

$$\begin{aligned} P(X/Y \le t) &= P(0 < Y < \infty, X/Y \le t) \\ &= \int_0^\infty \int_{-\infty}^{ty} f_X(x)\, f_Y(y)\, dxdy \\ &= \int_0^\infty \left[\int_{-\infty}^{ty} f_X(x)\, dx\right] f_Y(y)\, dy. \end{aligned}$$

Let $G(t) = \int_{-\infty}^{ty} f_X(x)\, dx$. We write

$$F_{X/Y}(t) = P(X/Y \le t) = \int_0^\infty G(t)\, f_Y(y)\, dy.$$

This implies

$$F'_{X/Y}(t) = \int_{-\infty}^\infty G'(t)\, f_Y(y)\, dy$$

i.e.,

$$f_{X/Y}(t) = \int_{-\infty}^\infty y f_X(ty)\, f_Y(y)\, dy$$

We see

$$
\begin{aligned}
f_{Z_1}(x) &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \\
f_{|Z_2|}(y) &= \frac{2}{\sqrt{2\pi}} e^{-y^2/2}
\end{aligned}
$$

Then

$$
\begin{aligned}
f_{Z_1/|Z_2|}(t) &= \frac{1}{\pi} \int_0^\infty y \exp\left(-\frac{(ty)^2}{2} - \frac{y^2}{2}\right) dy \\
&= \frac{1}{\pi} \int_0^\infty \left[-\frac{1}{1+t^2} \exp\left(-\frac{(ty)^2}{2} - \frac{y^2}{2}\right)\right]' dy = \frac{1}{\pi} \frac{1}{1+t^2}.
\end{aligned}
$$

Students' t distribution.

$$
T = \frac{Z_1}{\sqrt{\frac{Z_2^2 + \ldots + Z_n^2}{n-1}}} \sim t_{n-1}
$$

Similarly we have

$$
f_T(t) = \frac{\Gamma(n/2)}{\sqrt{(n-1)\pi}\,\Gamma\left(\frac{n-1}{2}\right)} \left(1 + \frac{t^2}{n-1}\right)^{-n/2}.
$$