Probability Theory Final preparation

0712238 林彥彤

https://scholar.harvard.edu/files/charlescywang/files/basic_statistics_and_probability_for_econometrics_econ_270a.pdf

Table of Common Distributions

taken from *Statistical Inference* by Casella and Berger

Discrete Distrbutions

distribution	pmf	mean	variance	mgf/moment
Bernoulli(p)	$p^x(1-p)^{1-x};\ x=0,1;\ p\in(0,1)$	p	$p(1-p)$	$(1-p)+pe^t$
Beta-binomial(n,α,β)	$\binom{n}{x}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(x+\alpha)\Gamma(n-x+\beta)}{\Gamma(\alpha+\beta+n)}$	$\frac{n\alpha}{\alpha+\beta}$	$\frac{n\alpha\beta}{(\alpha+\beta)^2}$	
Notes: If $X\|P$ is binomial (n,P) and P is beta(α,β), then X is beta-binomial(n,α,β).				
Binomial(n,p)	$\binom{n}{x}p^x(1-p)^{n-x};\ x=1,\ldots,n$	np	$np(1-p)$	$[(1-p)+pe^t]^n$
Discrete Uniform(N)	$\frac{1}{N};\ x=1,\ldots,N$	$\frac{N+1}{2}$	$\frac{(N+1)(N-1)}{12}$	$\frac{1}{N}\sum_{i=1}^{N}e^{it}$
Geometric(p)	$p(1-p)^{x-1};\ p\in(0,1)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$
Note: $Y=X-1$ is negative binomial$(1,p)$. The distribution is *memoryless*: $P(X>s\|X>t)=P(X>s-t)$.				
Hypergeometric(N,M,K)	$\frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}};\ x=1,\ldots,K$	$\frac{KM}{N}$	$\frac{KM}{N}\frac{(N-M)(N-k)}{N(N-1)}$	?
	$M-(N-K)\le x\le M;\ N,M,K>0$			
Negative Binomial(r,p)	$\binom{r+x-1}{x}p^r(1-p)^x;\ p\in(0,1)$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$	$\left(\frac{p}{1-(1-p)e^t}\right)^r$
	$\binom{y-1}{r-1}p^r(1-p)^{y-r};\ Y=X+r$			
Poisson(λ)	$\frac{e^{-\lambda}\lambda^x}{x!};\ \lambda\ge 0$	λ	λ	$e^{\lambda(e^t-1)}$
Notes: If Y is gamma(α,β), X is Poisson$(\frac{x}{\beta})$, and α is an integer, then $P(X\ge\alpha)=P(Y\le y)$.				

Continuous Distributions

distribution	pdf	mean	variance	mgf/moment
Beta(α,β)	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1};\ x\in(0,1),\ \alpha,\beta>0$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	$1+\sum_{k=1}^{\infty}\left(\prod_{r=0}^{k-1}\frac{\alpha+r}{\alpha+\beta+r}\right)\frac{t^k}{k!}$
Cauchy(θ,σ)	$\frac{1}{\pi\sigma}\frac{1}{1+(\frac{x-\theta}{\sigma})^2};\ \sigma>0$	does not exist	does not exist	does not exist
Notes: Special case of Students's t with 1 degree of freedom. Also, if X,Y are iid $N(0,1)$, $\frac{X}{Y}$ is Cauchy				
χ^2_p	$\frac{1}{\Gamma(\frac{p}{2})2^{\frac{p}{2}}}x^{\frac{p}{2}-1}e^{-\frac{x}{2}};\ x>0,\ p\in N$	p	$2p$	$\left(\frac{1}{1-2t}\right)^{\frac{p}{2}},\ t<\frac{1}{2}$
Notes: Gamma$(\frac{p}{2},2)$.				
Double Exponential(μ,σ)	$\frac{1}{2\sigma}e^{-\frac{\lvert x-\mu\rvert}{\sigma}};\ \sigma>0$	μ	$2\sigma^2$	$\frac{e^{\mu t}}{1-(\sigma t)^2}$
Exponential(θ)	$\frac{1}{\theta}e^{-\frac{x}{\theta}};\ x\geq 0,\ \theta>0$	θ	θ^2	$\frac{1}{1-\theta t},\ t<\frac{1}{\theta}$
Notes: Gamma$(1,\theta)$. Memoryless. $Y=X^{\frac{1}{\gamma}}$ is *Weibull*. $Y=\sqrt{\frac{2X}{\beta}}$ is *Rayleigh*. $Y=\alpha-\gamma\log\frac{X}{\beta}$ is *Gumbel*.				
F_{ν_1,ν_2}	$\frac{\Gamma(\frac{\nu_1+\nu_2}{2})}{\Gamma(\frac{\nu_1}{2})\Gamma(\frac{\nu_2}{2})}\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}\frac{x^{\frac{\nu_1-2}{2}}}{\left(1+(\frac{\nu_1}{\nu_2})x\right)^{\frac{\nu_1+\nu_2}{2}}};\ x>0$	$\frac{\nu_2}{\nu_2-2},\ \nu_2>2$	$2(\frac{\nu_2}{\nu_2-2})^2\frac{\nu_1+\nu_2-2}{\nu_1(\nu_2-4)},\ \nu_2>4$	$EX^n=\frac{\Gamma(\frac{\nu_1+2n}{2})\Gamma(\frac{\nu_2-2n}{2})}{\Gamma(\frac{\nu_1}{2})\Gamma(\frac{\nu_2}{2})}\left(\frac{\nu_2}{\nu_1}\right)^n,\ n<\frac{\nu_2}{2}$
Notes: $F_{\nu_1,\nu_2}=\frac{\chi^2_{\nu_1}/\nu_1}{\chi^2_{\nu_2}/\nu_2}$, where the χ^2s are independent. $F_{1,\nu}=t^2_\nu$.				
Gamma(α,β)	$\frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}e^{-\frac{x}{\beta}};\ x>0,\ \alpha,\beta>0$	$\alpha\beta$	$\alpha\beta^2$	$\left(\frac{1}{1-\beta t}\right)^{\alpha},\ t<\frac{1}{\beta}$
Notes: Some special cases are exponential $(\alpha=1)$ and χ^2 $(\alpha=\frac{p}{2},\beta=2)$. If $\alpha=\frac{2}{3}$, $Y=\sqrt{\frac{X}{\beta}}$ is *Maxwell*. $Y=\frac{1}{X}$ is *inverted gamma*.				
Logistic(μ,β)	$\frac{1}{\beta}\frac{e^{-\frac{x-\mu}{\beta}}}{\left[1+e^{-\frac{x-\mu}{\beta}}\right]^2};\ \beta>0$	μ	$\frac{\pi^2\beta^2}{3}$	$e^{\mu t}\Gamma(1+\beta t),\ \lvert t\rvert<\frac{1}{\beta}$
Notes: The cdf is $F(x\mid\mu,\beta)=\frac{1}{1+e^{-\frac{x-\mu}{\beta}}}$.				
Lognormal(μ,σ^2)	$\frac{1}{\sqrt{2\pi}\sigma}\frac{1}{x}e^{-\frac{(\log x-\mu)^2}{2\sigma^2}};\ x>0,\sigma>0$	$e^{\mu+\frac{\sigma^2}{2}}$	$e^{2(\mu+\sigma^2)}-e^{2\mu+\sigma^2}$	$EX^n=e^{n\mu+\frac{n^2\sigma^2}{2}}$
Normal(μ,σ^2)	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}};\ \sigma>0$	μ	σ^2	$e^{\mu t+\frac{\sigma^2t^2}{2}}$
Pareto(α,β)	$\frac{\beta\alpha^\beta}{x^{\beta+1}};\ x>\alpha,\ \alpha,\beta>0$	$\frac{\beta\alpha}{\beta-1},\ \beta>1$	$\frac{\beta\alpha^2}{(\beta-1)^2(\beta-2)},\ \beta>2$	does not exist
t_ν	$\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}\frac{1}{\sqrt{\nu\pi}}\frac{1}{(1+\frac{x^2}{\nu})^{\frac{\nu+1}{2}}}$	$0,\ \nu>1$	$\frac{\nu}{\nu-2},\nu>2$	$EX^n=\frac{\Gamma(\frac{\nu+1}{2})\Gamma(\nu-\frac{n}{2})}{\sqrt{\pi}\Gamma(\frac{\nu}{2})}\nu^{\frac{n}{2}}$, n even
Notes: $t^2_\nu=F_{1,\nu}$.				
Uniform(a,b)	$\frac{1}{b-a},\ a\leq x\leq b$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt}-e^{at}}{(b-a)t}$
Notes: If $a=0,\ b=1$, this is special case of beta $(\alpha=\beta=1)$.				
Weibull(γ,β)	$\frac{\gamma}{\beta}x^{\gamma-1}e^{-\frac{x^\gamma}{\beta}};\ x>0,\ \gamma,\beta>0$	$\beta^{\frac{1}{\gamma}}\Gamma(1+\frac{1}{\gamma})$	$\beta^{\frac{2}{\gamma}}\left[\Gamma(1+\frac{2}{\gamma})-\Gamma^2(1+\frac{1}{\gamma})\right]$	$EX^n=\beta^{\frac{n}{\gamma}}\Gamma(1+\frac{n}{\gamma})$
Notes: The mgf only exists for $\gamma\geq 1$.				

5.5 Geometric random variable

A Geometric(p) random variable X counts the number trials required for the first success in independent trials with success probability p.

Properties:

1. Probability mass function: $P(X=n)=p(1-p)^{n-1}$, where $n=1,2,\ldots$.
2. $EX=\frac{1}{p}$.
3. $\mathrm{Var}(X)=\frac{1-p}{p^2}$.
4. $P(X>n)=\sum_{k=n+1}^{\infty}p(1-p)^{k-1}=(1-p)^n$.
5. $P(X>n+k\mid X>k)=\frac{(1-p)^{n+k}}{(1-p)^k}=P(X>n)$.

5.4 Poisson random variable

A random variable is Poisson(λ), with parameter $\lambda > 0$, if it has the probability mass function given below.

Properties:

1. $P(X = i) = \frac{\lambda^i}{i!} e^{-\lambda}$, for $i = 0, 1, 2, \ldots$.
2. $EX = \lambda$.
3. $\mathrm{Var}(X) = \lambda$.

Here is how we compute the expectation:

$$EX = \sum_{i=1}^{\infty} i \cdot e^{-\lambda} \frac{\lambda^i}{i!} = e^{-\lambda} \lambda \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} = e^{-\lambda} \lambda \, e^{\lambda} = \lambda,$$

and the variance computation is similar (and a good exercise!).

6.2 Exponential random variable

A random variable is Exponential(λ), with parameter $\lambda > 0$, if it has the probability mass function given below. This is a distribution for the *waiting time* for some random event, for example, for a lightbulb to burn out or for the next earthquake of at least some given magnitude.

Properties:

1. Density: $f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases}$
2. $EX = \frac{1}{\lambda}$.
3. $\mathrm{Var}(X) = \frac{1}{\lambda^2}$.
4. $P(X \geq x) = e^{-\lambda x}$.
5. Memoryless property: $P(X \geq x + y | X \geq y) = e^{-\lambda x}$.

Conditional distributions

The conditional p. m. f. of X given $Y = y$ is, in the discrete case, given simply by

$$p_X(x|Y=y) = P(X=x|Y=y) = \frac{P(X=x, Y=y)}{P(Y=y)}.$$

This is trickier in the continuous case, as we cannot divide by $P(Y = y) = 0$.

For a jointly continuous pair of random variables X and Y, we define the *conditional density of X given $Y = y$* as follows:

$$f_X(x|Y=y) = \frac{f(x,y)}{f_Y(y)},$$

where $f(x, y)$ is, of course, the joint density of (X, Y).

$$E(\alpha_1 X_1 + \alpha_2 X_2) = \alpha_1 EX_1 + \alpha_2 EX_2,$$

om variables X_1 and X_2 and nonrandom constants α_1 and α_2. This proper
and discussed in more detail later. Then

$$\begin{aligned}
\mathrm{Var}(X) &= E[(X-\mu)^2] \\
&= E[X^2 - 2\mu X + \mu^2] \\
&= E(X^2) - 2\mu E(X) + \mu^2 \\
&= E(X^2) - \mu^2 = E(X^2) - (EX)^2
\end{aligned}$$

To summarize, the most useful formula is

$$\boxed{\mathrm{Cov}(X,Y) = E(XY) - EX \cdot EY.}$$

Note immediately that, if X and Y are independent, then $\mathrm{Cov}(X, Y) = 0$, but the converse is false.

Let X and Y be indicator random variables, so $X = I_A$ and $Y = I_B$, for two events A and B. Then, $EX = P(A)$, $EY = P(B)$, $E(XY) = E(I_{A\cap B}) = P(A \cap B)$, and so

$$\mathrm{Cov}(X,Y) = P(A\cap B) - P(A)P(B) = P(A)[P(B|A) - P(B)].$$

If $P(B|A) > P(B)$, we say the two events are *positively correlated* and, in this case, the covariance is positive; if the events are *negatively correlated* all inequalities are reversed. For general random variables X and Y, $\mathrm{Cov}(X, Y) > 0$ intuitively means that, "on the average," increasing X will result in larger Y.

Variance of sums of random variables

Theorem 8.4. *Variance-covariance formula:*

$$\boxed{\begin{aligned}
E(\sum_{i=1}^n X_i)^2 &= \sum_{i=1}^n EX_i^2 + \sum_{i\neq j} E(X_i X_j), \\
\mathrm{Var}(\sum_{i=1}^n X_i) &= \sum_{i=1}^n \mathrm{Var}(X_i) + \sum_{i\neq j} \mathrm{Cov}(X_i, X_j).
\end{aligned}}$$

1.3 變異數性質

1. 變異量數不會是負的
2. 當一個資料集的變異量數為零時，其內所有項目皆為相同數值
3. 一個常數被加至一個數列中的所有變數值，此數列的變異量數不會改變：$\operatorname{Var}(X+a)=\operatorname{Var}(X.)$
4. 如果所有數值被放大一個常數倍，變異量數會放大此常數的次方倍：$\operatorname{Var}(aX)=a^2\operatorname{Var}(X)$
5. 兩個隨機變數合的變異量數為：$\operatorname{Var}(aX+bY)=a^2\operatorname{Var}(X)+b^2\operatorname{Var}(Y)+2ab\ \operatorname{Cov}(X,Y),$
6. 對於N個隨機變數$\{X_1,\dots,X_N\}$的總和：
   $\operatorname{Var}\left(\sum_{i=1}^N X_i\right)=\sum_{i,j=1}^N \operatorname{Cov}(X_i,X_j)=\sum_{i=1}^N \operatorname{Var}(X_i)+\sum_{i\neq j}\operatorname{Cov}(X_i,X_j)$

$$\mu_{a+bX}=a+b\mu_X \qquad \sigma^2_{a+bX}=b^2\sigma^2$$

$$\sigma^2_{X+Y}=\sigma^2_X+\sigma^2_Y$$

$$\sigma^2_{X-Y}=\sigma^2_X+\sigma^2_Y$$

FOR INDEPENDENT X Y(Cov(X,Y)=E(XY)-E(X)*E(Y) = 0)

Theorem 8.7. *Markov Inequality. If $X \geq 0$ is a random variable and $a > 0$, then*

$$P(X \geq a) \leq \frac{1}{a} EX.$$

Example 8.15. If $EX = 1$ and $X \geq 0$, it must be that $P(X \geq 10) \leq 0.1$.

Proof. Here is the crucial observation:

$$I_{\{X \geq a\}} \leq \frac{1}{a} X.$$

Indeed, if $X < a$, the left-hand side is 0 and the right-hand side is nonnegative; if $X \geq a$, the left-hand side is 1 and the right-hand side is at least 1. Taking the expectation of both sides, we get

$$P(X \geq a) = E(I_{\{X \geq a\}}) \leq \frac{1}{a} EX.$$

□

Theorem 8.8. *Chebyshev inequality. If $EX = \mu$ and $\mathrm{Var}(X) = \sigma^2$ are both finite and $k > 0$, then*

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}.$$

Example 8.16. If $EX = 1$ and $\mathrm{Var}(X) = 1$, $P(X \geq 10) \leq P(|X - 1| \geq 9) \leq \frac{1}{81}$.

Example 8.17. If $EX = 1$, $\mathrm{Var}(X) = 0.1$,

$$P(|X - 1| \geq 0.5) \leq \frac{0.1}{0.5^2} = \frac{2}{5}.$$

If $X, X_1, X_2, \ldots$ are independent and identically distributed random variables with finite expectation and variance, then $\frac{X_1 + \ldots + X_n}{n}$ converges to EX in the sense that, for any fixed $\epsilon > 0$,

$$P\left(\left|\frac{X_1 + \ldots + X_n}{n} - EX\right| \geq \epsilon\right) \to 0,$$

as $n \to \infty$.

Central Limit Theorem

Theorem 8.9. *Central limit theorem.*

Assume that $X, X_1, X_2, \ldots$ are independent, identically distributed random variables, with finite $\mu = EX$ and $\sigma^2 = Var(X)$. Then,

$$P\left(\frac{X_1 + \ldots + X_n - \mu n}{\sigma \sqrt{n}} \leq x\right) \to P(Z \leq x),$$

as $n \to \infty$, where Z is standard Normal.

6. Let $X_1, X_2, \ldots$ be the numbers on successive rolls and $S_n = X_1 + \ldots + X_n$ the sum. We know that $EX_i = \frac{7}{2}$, and $\mathrm{Var}(X_i) = \frac{35}{12}$. So, we have

$$P(S_{24} \geq 100) = P\left(\frac{S_{24} - 24 \cdot \frac{7}{2}}{\sqrt{24 \cdot \frac{35}{12}}} \geq \frac{100 - 24 \cdot \frac{7}{2}}{\sqrt{24 \cdot \frac{35}{12}}}\right) \approx P(Z \geq 1.85) = 1 - \Phi(1.85) \approx 0.032.$$

Theorem 9.1. *Connection between variance and convergence in probability.*

Assume that Y_n are random variables and that a is a constant such that

$$EY_n \to a,$$
$$\mathrm{Var}(Y_n) \to 0,$$

as $n \to \infty$. Then,

$$Y_n \to a,$$

as $n \to \infty$, in probability.

10 Moment generating functions

If X is a random variable, then its *moment generating function* is

$$\phi(t) = \phi_X(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx} P(X = x) & \text{in the discrete case,} \\ \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx & \text{in the continuous case.} \end{cases}$$

$$\begin{aligned} E(e^{tX}) &= E[1 + tX + \frac{1}{2}t^2X^2 + \frac{1}{3!}t^3X^3 + \ldots] \\ &= 1 + tE(X) + \frac{1}{2}t^2E(X^2) + \frac{1}{3!}t^3E(X^3) + \ldots \end{aligned}$$

Moment Generating Functions: $M_x(t) = E\left(e^{tX}\right)$ and $E\left[X^{(n)}\right] = M_X^{(n)}(0)$ *where* $M_X^{(n)} = \frac{\partial^{(n)}}{\partial t} M_X(t)$

- Useful Properties of MGF: If X,Y **independent**
  $M_{aX+b}(t) = \exp(bt) M_X(at)$
  $M_{X+Y}(t) = M_X(t) M_Y(t)$

MGF of a Sample Average (of a random sample): $M_{\overline{X}}(t) = M_{\frac{1}{N}\left(\sum X_i\right)}(t) = \prod M_X\left(\frac{t}{N}\right)$

USE OF INDICATOR!!!!

USE OF RECURSION!!!!

11 Computing probabilities and expectations by conditioning

Conditioning is the method we encountered before; to remind ourselves, it involves two-stage (or multistage) processes and conditions are appropriate events at the first stage. Recall also the basic definitions:

- *Conditional probability*: if A and B are two events, $P(A|B) = \frac{P(A \cap B)}{P(B)}$;
- *Conditional probability mass function*: if (X, Y) has probability mass function p, $p_X(x|Y = y) = \frac{p(x,y)}{p_Y(y)} = P(X = x|Y = y)$;
- *Conditional density*: if (X, Y) has joint density f, $f_X(x|Y = y) = \frac{f(x,y)}{f_Y(y)}$.
- *Conditional expectation*: $E(X|Y = y)$ is either $\sum_x x p_X(x|Y = y)$ or $\int x f_X(x|Y = y)\, dx$ depending on whether the pair (X, Y) is discrete or continuous.

Bayes' formula also applies to expectation. Assume that the distribution of a random variable X conditioned on $Y = y$ is *given*, and, consequently, its expectation $E(X|Y = y)$ is also known. Such is the case of a two-stage process, whereby the value of Y is chosen at the first stage, which then determines the distribution of X at the second stage. This situation is very common in applications. Then,

$$\boxed{E(X) = \begin{cases} \sum_y E(X|Y = y)P(Y = y) & \text{if } Y \text{ is discrete,} \\ \int_{-\infty}^{\infty} E(X|Y = y) f_Y(y)\, dy & \text{if } Y \text{ is continuous.} \end{cases}}$$

Note that this applies to the probability of an event (which is nothing other than the expectation of its indicator) as well — if we know $P(A|Y = y) = E(I_A|Y = y)$, then we may compute $P(A) = EI_A$ by Bayes' formula above.

Example 8.6. *Coupon collector problem*, revisited. Sample from n cards, with replacement, indefinitely. Let N be the number of cards you need to sample for a complete collection, i.e., to get all different cards represented. What is EN?

Let N_i be the number of *additional* cards you need to get the ith *new* card, *after* you have received the $(i-1)$st new card.

Then, N_1, the number of cards needed to receive the first new card, is trivial, as the first card you buy is new: $N_1 = 1$. Afterward, N_2, the number of additional cards needed to get the second new card is Geometric with success probability $\frac{n-1}{n}$. After that, N_3, the number of additional cards needed to get the third new card is Geometric with success probability $\frac{n-2}{n}$. In general, N_i is geometric with success probability $\frac{n-i+1}{n}$, $i = 1, \ldots, n$, and

$$N = N_1 + \ldots + N_n,$$

so that

$$EN = n\left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}\right).$$

Now, we have

$$\sum_{i=2}^{n} \frac{1}{i} \le \int_1^n \frac{1}{x}\,dx \le \sum_{i=1}^{n-1} \frac{1}{i},$$

by comparing the integral with the Riemman sum at the left and right endpoints in the division of $[1,n]$ into $[1,2]$, $[2,3]$, $\ldots$, $[n-1,n]$, and so

$$\log n \le \sum_{i=1}^{n} \frac{1}{i} \le \log n + 1,$$

which establishes the limit

$$\lim_{n\to\infty} \frac{EN}{n \log n} = 1.$$

Example 11.2. Let T_1, T_2 be two independent Exponential(λ) random variables and let $S_1 = T_1$, $S_2 = T_1 + T_2$. Compute $f_{S_1}(s_1 | S_2 = s_2)$.

First,

$$\begin{aligned} P(S_1 \le s_1, S_2 \le s_2) &= P(T_1 \le s_1, T_1 + T_2 \le s_2) \\ &= \int_0^{s_1} dt_1 \int_0^{s_2 - t_1} f_{T_1,T_2}(t_1, t_2)\, dt_2. \end{aligned}$$

If $f = f_{S_1,S_2}$, then

$$\begin{aligned} f(s_1, s_2) &= \frac{\partial^2}{\partial s_1 \partial s_2} P(S_1 \le s_1, S_2 \le s_2) \\ &= \frac{\partial}{\partial s_2} \int_0^{s_2 - s_1} f_{T_1,T_2}(s_1, t_2)\, dt_2 \\ &= f_{T_1,T_2}(s_1, s_2 - s_1) \\ &= f_{T_1}(s_1) f_{T_2}(s_2 - s_1) \\ &= \lambda e^{-\lambda s_1} \lambda\, e^{-\lambda(s_2 - s_1)} \\ &= \lambda^2 e^{-\lambda s_2}. \end{aligned}$$

Therefore,

$$f(s_1, s_2) = \begin{cases} \lambda^2 e^{-\lambda s_2} & \text{if } 0 \le s_1 \le s_2, \\ 0 & \text{otherwise} \end{cases}$$

and, consequently, for $s_2 \ge 0$,

$$f_{S_2}(s_2) = \int_0^{s_2} f(s_1, s_2)\, ds_1 = \lambda^2 s_2\, e^{-\lambda s_2}.$$

Therefore,

$$f_{S_1}(s_1 | S_2 = s_2) = \frac{\lambda^2 e^{-\lambda s_2}}{\lambda^2 s_2\, e^{-\lambda s_2}} = \frac{1}{s_2},$$

for $0 \le s_1 \le s_2$, and 0 otherwise. Therefore, conditioned on $T_1 + T_2 = s_2$, T_1 is uniform on $[0, s_2]$.

Imagine the following: a new lightbulb is put in and, after time T_1, it burns out. It is then replaced by a new lightbulb, identical to the first one, which also burns out after an additional time T_2. If we know the time when the second bulb burns out, the first bulb's failure time is uniform on the interval of its possible values.

Example 11.7. *Bold Play.* Assume that the only game available to you is a game in which you can place even bets at any amount, and that you win each of these bets with probability p. Your initial capital is $x \in [0, N]$, a real number, and again you want to increase it to N before going broke. Your bold strategy (which can be proved to be the best) is to bet everything unless you are close enough to N that a smaller amount will do:

1. Bet x if $x \leq \frac{N}{2}$.
2. Bet $N - x$ if $x \geq \frac{N}{2}$.

We can, without loss of generality, fix our monetary unit so that $N = 1$. We now define

$$P(x) = P(\text{reach 1 before reaching 0}).$$

By conditioning on the outcome of your first bet,

$$P(x) = \begin{cases} p \cdot P(2x) & \text{if } x \in \left[0, \frac{1}{2}\right], \\ p \cdot 1 + (1-p) \cdot P(2x-1) & \text{if } x \in \left[\frac{1}{2}, 1\right]. \end{cases}$$

For each positive integer n, this is a linear system for $P(\frac{k}{2^n})$, $k = 0, \ldots, 2^n$, which can be solved. For example:

- When $n = 1$, $P\left(\frac{1}{2}\right) = p$.
- When $n = 2$, $P\left(\frac{1}{4}\right) = p^2$, $P\left(\frac{3}{4}\right) = p + (1-p)p$.
- When $n = 3$, $P\left(\frac{1}{8}\right) = p^3$, $P\left(\frac{3}{8}\right) = p \cdot P\left(\frac{3}{4}\right) = p^2 + p^2(1-p)$, $P\left(\frac{5}{8}\right) = p + p^2(1-p)$, $P\left(\frac{7}{8}\right) = p + p(1-p) + p(1-p)^2$.

It is easy to verify that $P(x) = x$, for all x, if $p = \frac{1}{2}$. Moreover, it can be computed that $P(0.9) \approx 0.8794$ for $p = \frac{9}{19}$, which is not too different from a fair game. The figure below displays the graphs of functions $P(x)$ for $p = 0.1$, 0.25, $\frac{9}{19}$, and $\frac{1}{2}$.

Assume that $X, X_1, X_2, \ldots$ is an i. i. d. sequence of random variables with finite $EX = \mu$ and $\text{Var}(X) = \sigma^2$. Let N be a nonnegative integer random variable, independent of all X_i, and let

$$S = \sum_{i=1}^{N} X_i.$$

Then

$$ES = \mu\, EN,$$
$$\text{Var}(S) = \sigma^2\, EN + \mu^2 \text{Var}(N).$$

MARKOV 12 IS TRIVIAL
COUPON x1+x2+...xk<s Ganble Bold Bet etc0

.

Gambler's ruin $px_{j+1} - (p+q)x_j + qx_{j-1} \;=\; 0\,, \quad x_0 = 0, x_N = 1$

With $x_j = \alpha^j$ we find the quadratic equation

$$p\alpha^2 - (p+q)\alpha + q = 0$$

with solutions

$$\alpha \;=\; \frac{-p \pm \sqrt{(p+q)^2 - 4pq}}{2p} \;=\; \frac{-p \pm \sqrt{p^2 + q^2 - 2pq}}{2p} \;=\; \frac{-p \pm \sqrt{(p-q)^2}}{2p} \;=\; \left\{ \begin{array}{c} 1 \\ q/p \end{array} \right.$$

If $p \neq$ we have two solutions and and so the general solution is given by

$$x_n \;=\; C_1 1^n + C_2 \left(\frac{q}{p}\right)^n$$

...